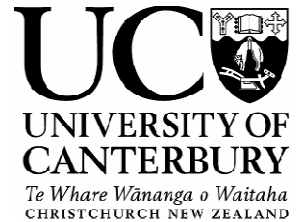


31 July 2007



---

The Effect of Road Network Bendiness on Traffic Crash  
Occurrence in New Zealand

A thesis submitted in partial fulfilment of the requirements  
for the Degree of

Master of Engineering in Transportation

in the University of Canterbury

by

Megan Fowler

## **Abstract**

Recent researchers have suggested that the combination of horizontal curves or “bendiness” of a length of road contributes to traffic crash occurrence. A previous study of New Zealand fatal crashes using an aggregated approach found no significant correlation between crash occurrence and road bendiness for rural roads but a minor correlation for urban roads.

This thesis further explores the effect of road bendiness on traffic crash occurrence in New Zealand by developing a method more suited to traffic engineering. The method involves Geographical Information Systems (GIS) firstly to process data and secondly to calculate bendiness values. The following bendiness measures: bend density, detour ratio, cumulative angle, mean angle and standard deviation of angles; are applied to “influence areas” surrounding crash and comparison sites. The method then dictates that some form of statistical analysis should be performed to distinguish between the bendiness of crash and comparison sites, while accounting for other influencing factors. Binary logistic regression is recommended.

The method was applied in a case study of New Zealand fatal crashes, with two main analysis techniques employed. Firstly, binary logistic regression models were developed. It was found that, for rural roads, sections with consistent and frequent curves were safer than completely straight sections or those with isolated curves. The urban model was less conclusive, which suggests that the method was not appropriate in the urban situation.

The second analysis method involved comparing bendiness values of a site’s “immediate area” with those of its influence area. It was found that, although the spreads of the comparison sites’ distributions were smaller than those of the crash sites, the mean values were generally very similar and no appropriate bendiness ratios could be specified to reduce crash risk.

Overall it appears that, if design consistency is maintained, bendiness is a protective quality for rural roads.

## **Dedications**

A million thanks to Mum, Dad, Roslyn, Bridget, Jacob, Kate, Hannah, Sarah and David for all your love and support throughout my studies and life in general.

My strength in completing this thesis and all other areas of my life comes from Jesus Christ and I offer to Him all that comes from the gifts He entrusts to me.

*“Look to Him in every course you take, and He will see that your roads are smooth.”*

*– Proverbs 3:6*

Also, there have been a large number of people who have assisted me with this thesis whom I am hugely indebted to and wish to acknowledge and thank in Section 9.0.

## Table of Contents

List of Figures .....	v
List of Tables .....	vii
Glossary .....	viii
1.0 Introduction .....	1
2.0 Literature Review .....	4
2.1 General Background Literature .....	4
2.1.1 Curves as Contributing Crash Factors .....	4
2.1.2 Bendiness .....	7
2.1.3 Design Consistency.....	11
2.1.4 Driver Awareness of the Road Environment .....	16
2.1.5 Demand of the Road Environment on Drivers .....	18
2.1.6 Use of GIS Techniques .....	19
2.1.7 Level of Aggregation .....	21
2.1.8 Statistical Evaluation of Results .....	22
2.2 Motivating Studies.....	24
2.2.1 District variations in road curvature in England and Wales and their association with road traffic crashes (Haynes <i>et al.</i> 2007a) .....	24
2.2.2 Influence of Road Curvature on Fatal Crashes in New Zealand (Haynes <i>et al.</i> 2007b) .....	29
2.3 Implications for this Research.....	31
2.3.1 Variables to be Considered .....	31
2.3.2 Methods of Evaluation .....	32
2.3.3 Suggested Improvements to Previous Study Methods .....	33
3.0 General Method.....	37
3.1 Data Preparation .....	37
3.2 Calculations of Bendiness Measures .....	42
3.2.1 Method 1: A simple summation of each route's bendiness values.....	44
3.2.2 Method 2: Weighted according to the number of routes in the area .....	45
3.2.3 Method 3: Weighted according to number of routes at each intersection ..	45
3.2.4 Method 4: Weighted according to routes' link flows .....	46
3.3 Analysis Methods .....	46
3.3.1 Comparison Data .....	46
3.3.2 Choice of Statistical Model .....	47
4.0 Case Study .....	49
4.1 General Description of New Zealand .....	49
4.2 Data Relating to Bendiness .....	50
4.2.1 Description of Crash Data .....	51
4.2.2 Description of Road Network Data .....	56
4.2.3 Modifications made to Crash Data .....	58
4.2.4 Modifications made to Road Network Data.....	60
4.3 Comparison Locations .....	65
4.4 Data Relating to Possible Other Influencing Variables.....	68
4.4.1 Upper Hutt City Flow Data .....	69
4.4.2 State Highway Flow Data .....	72
4.4.3 Elevation and Rainfall Data .....	73
4.4.4 Employment and Population Data .....	75
4.4.5 Junction Density .....	75
4.4.6 Intersection versus Mid-block Locations .....	76
4.4.7 Variables Intentionally Excluded from Investigation .....	77

4.4.8 Desired but Unattained Data .....	78
5.0 Case Study Bendiness Measure Calculations .....	79
5.1 Applicability to Case Study of Bendiness Measures used in Motivating Study	79
5.2 Calculation of Bendiness Measures used in the Motivating Study .....	83
5.3 Additional Bendiness Measures created for Case Study .....	92
5.3.1 Standard Deviation between Angles.....	92
5.3.2 Comparison of Measures for Immediate Area versus Influence Area.....	93
6.0 Preliminary Case Study Investigations .....	95
6.1 Comparisons with Motivating Study .....	95
6.1.1 Rural Road Crashes .....	95
6.1.2 Urban Road Crashes .....	102
6.2 Non-Flow Weighted Comparisons between Crash and Non-Crash Sites .....	106
6.2.1 Rural Roads .....	106
6.2.2 Urban Roads .....	112
6.3 Bendiness Measures applied with Flow Weighting in Upper Hutt Region .....	116
7.0 Case Study Results .....	121
7.1 Model Estimation .....	121
7.1.1 Rural State Highway Crash Model .....	121
7.1.2 Rural non-State Highway Crash Model .....	129
7.1.3 Urban Crash Model.....	135
7.1.4 Comparative Analysis of Models .....	141
7.2 Immediate versus Influence Area Bendiness Measure Evaluations.....	143
8.0 Conclusions .....	148
8.1 Summary of Findings .....	148
8.1.1 Initial Comparisons with Motivating Study's Measures .....	148
8.1.2 Crash versus Non-Crash Comparisons of Bendiness Measures.....	149
8.1.3 Use of Flow Weighting Techniques for Bendiness Measure Calculation	149
8.1.4 Binary Logistic Regression Models.....	150
8.1.5 Immediate versus Influence Area Analysis.....	151
8.2 Implications of Findings with respect to Objectives .....	151
8.3 Problems Encountered.....	153
8.3.1 Computer software.....	153
8.3.2 Data formats .....	154
8.3.3 Selection of comparison sites .....	154
8.4 Areas for Future Research .....	155
8.4.1 Investigation of influence area sizes.....	155
8.4.2 Alternative weighting methods.....	156
8.4.3 Determination of actual routes travelled.....	157
8.4.4 Incorporate the effects of hilliness.....	158
8.4.5 Include injury crashes in analysis.....	158
8.4.6 Design alternative measures to represent design consistency .....	159
8.4.7 Relate findings to knowledge of rate of crashes per individual curve.....	160
9.0 Acknowledgements.....	162
10.0 References .....	164
11.0 Appendix .....	169
11.1 Appendix 1: Summary of Crash Data.....	169
11.2 Appendix 2: Model Correlations .....	170
11.2.1 Rural State Highways.....	170
11.2.2 Rural non-State Highways .....	172
11.2.3 Urban Roads .....	174

## List of Figures

Figure 2.1 Two Types of Horizontal Curvature (adapted from Vis (2007)) .....	4
Figure 2.2 Plan View of Two "Bendy" Roads.....	7
Figure 2.3 Plan View of Two Sample Roads to Illustrate Design Consistency .....	11
Figure 2.4 Chevron Curve Warning Sign (from Transit New Zealand <i>et al.</i> (2007)) .	17
Figure 2.5 Herringbone Pavement Marking (from Charlton (2007)) .....	17
Figure 2.6 Components of Road Network Data (adapted from Haynes <i>et al.</i> (2007a)) .....	25
Figure 2.7 Bendiness Diagram (adapted from Haynes <i>et al.</i> (2007a)).....	25
Figure 2.8 Concept Showing Potential Pitfalls of Aggregation .....	34
Figure 3.1 Data Preparation Flow Diagram.....	39
Figure 3.2 Processed Road Network Concept Diagram.....	42
Figure 3.3 Bendiness Measures Calculations Flow Diagram.....	43
Figure 3.4 Data for Statistical Analysis.....	47
Figure 4.1 Map of New Zealand (adapted from LINZ (2005)) .....	49
Figure 4.2 Bend Density Frequency for different CAS Curve Types.....	53
Figure 4.3 Cumulative Angle Frequency for different CAS Curve Types .....	54
Figure 4.4 Cumulative Frequency of Speed Limit versus Fatalities.....	56
Figure 4.5 Digitised Road Network Compared to Satellite Image (from (Google Earth 2007)) .....	58
Figure 4.6 Cumulative Distribution of Distance between Crashes and Road .....	59
Figure 4.7 Example of Arc Division .....	60
Figure 4.8 Example of Link Division at Crash Location .....	61
Figure 4.9 Part of the Processed Network .....	63
Figure 4.10 Duplicate Link in Original Data.....	64
Figure 4.11 Rural and Urban Comparison Vertices.....	67
Figure 4.12 Upper Hutt City Location .....	70
Figure 4.13 Comparison of New Zealand and Upper Hutt Road Networks .....	71
Figure 4.14 Road Types for Rural Crash 1km Influence Area.....	73
Figure 4.15 Road Types for Rural Comparison Location 1km Influence Area .....	73
Figure 4.16 North and South Island Digital Elevation Models .....	74
Figure 4.17 Mean Precipitation Model for New Zealand .....	74
Figure 4.18 Proportion of Intersection Sites in Crash and Comparison Samples.....	77
Figure 5.1 Approximation of Circular Curve with Straight Sections .....	80
Figure 5.2 Effect of Resolution on Bend Density Ratio .....	80
Figure 5.3 Effect of Resolution on Detour Ratio.....	81
Figure 5.4 Effect of Resolution on Cumulative Angle Ratio .....	81
Figure 5.5 Effect of Resolution on Mean Angle Ratio .....	82
Figure 5.6 1km Influence Area around Crash .....	84
Figure 5.7 Identification of External Vertices .....	86
Figure 5.8 Example Route .....	87
Figure 5.9 Incorrect Route without using Barriers .....	88
Figure 5.10 Correct Route using Barriers.....	88
Figure 5.11 Urban Situation where Barriers are Unwanted .....	89
Figure 5.12 Rural Situation where Barriers are Unwanted .....	90
Figure 5.13 Ordered Route .....	91
Figure 5.14 Immediate and Influence Areas for Crash 2200072.....	93
Figure 6.1 Bend Density Comparison by TLAs for Rural State Highways .....	96
Figure 6.2 Bend Density Comparison by TLAs for Rural non-State Highway Roads	96
Figure 6.3 Detour Ratio Comparison by TLAs for Rural State Highways .....	97

Figure 6.4 Detour Ratio Comparison by TLAs for Rural non-State Highway Roads .	97
Figure 6.5 Cumulative Angle Comparison by TLAs for Rural State Highways .....	98
Figure 6.6 Cumulative Angle Comparison by TLAs for Rural non-State Highway Roads.....	98
Figure 6.7 Mean Angle Comparison by TLAs for Rural State Highways .....	99
Figure 6.8 Mean Angle Comparison by TLAs for Rural non-State Highways .....	99
Figure 6.9 Cumulative Frequency Distributions of Percentage Differences between Haynes and Fowler measures for Rural Roads.....	100
Figure 6.10 Illustration of Nodes .....	102
Figure 6.11 Bend Density Comparison by TLAs for Urban Roads.....	103
Figure 6.12 Detour Ratio Comparison by TLAs for Urban Roads.....	103
Figure 6.13 Cumulative Angle Comparison by TLAs for Urban Roads .....	104
Figure 6.14 Mean Angle Comparison by TLAs for Urban Roads.....	104
Figure 6.15 Cumulative Frequency Distributions of Percentage Differences between Haynes and Fowler measures for Urban Roads .....	105
Figure 6.16 Bend Density for Rural State Highway Crash and Comparison Sites....	107
Figure 6.17 Detour Ratio for Rural State Highway Crash and Comparison Sites....	107
Figure 6.18 Cumulative Angle for Rural State Highway Crash and Comparison Sites .....	108
Figure 6.19 Mean Angle for Rural State Highway Crash and Comparison Sites .....	108
Figure 6.20 Standard Deviation of Angles for Rural State Highway Crash and Comparison Sites .....	109
Figure 6.21 Bend Density for Rural Non-State Highway Crash and Comparison Sites .....	110
Figure 6.22 Detour Ratio for Rural Non-State Highway Crash and Comparison Sites .....	110
Figure 6.23 Cumulative Angle for Rural Non-State Highway Crash and Comparison Sites.....	111
Figure 6.24 Mean Angle for Rural Non-State Highway Crash and Comparison Sites .....	111
Figure 6.25 Standard Deviation of Angles for Rural Non-State Highway Crash and Comparison Sites .....	112
Figure 6.26 Bend Density for Urban Crash and Comparison Sites .....	113
Figure 6.27 Detour Ratio for Urban Crash and Comparison Sites .....	113
Figure 6.28 Cumulative Angle for Urban Crash and Comparison Sites.....	114
Figure 6.29 Mean Angle for Urban Crash and Comparison Sites .....	114
Figure 6.30 Standard Deviation of Angles for Urban Crash and Comparison Sites .	115
Figure 6.31 Upper Hutt Crashes .....	116
Figure 6.32 Bend Density Results for Upper Hutt Crashes with and without Flow Weighting .....	117
Figure 6.33 Detour Ratio Results for Upper Hutt Crashes with and without Flow Weighting .....	118
Figure 6.34 Cumulative Angle Results for Upper Hutt Crashes with and without Flow Weighting .....	118
Figure 6.35 Mean Angle Results for Upper Hutt Crashes with and without Flow Weighting .....	119
Figure 6.36 Standard Deviation of Angles Results for Upper Hutt Crashes with and without Flow Weighting .....	119
Figure 7.1 Rural State Highway Model Crash Risk Predictions .....	126
Figure 7.2 Rural State Highway Crash Probabilities .....	127

Figure 7.3 Rural non-State Highway Model Crash Risk Predictions .....	133
Figure 7.4 Rural non-State Highway Crash Probabilities .....	134
Figure 7.5 Urban Roads Model Crash Risk Probabilities .....	139
Figure 7.6 Urban Typical Examples .....	140
Figure 7.7 Immediate versus Influence Area Bend Density for Rural Sites .....	144
Figure 7.8 Immediate versus Influence Area Detour Ratio for Rural Sites .....	144
Figure 7.9 Immediate versus Influence Area Cumulative Angle for Rural Sites .....	145
Figure 7.10 Immediate versus Influence Area Mean Angle for Rural Sites .....	145
Figure 7.11 Immediate versus Influence Area Standard Deviation of Angles for Rural Sites .....	146
Figure 8.1 Pre-processed Road Network Concept .....	160

## List of Tables

Table 4.1 Principal Data Sets and Sources .....	50
Table 4.2 New Zealand Crashes on Curves (1996-2005) .....	52
Table 4.3 Data Sets and Sources of Possible Influencing Variables .....	69
Table 7.1 Rural State Highway Model .....	122
Table 7.2 Summary of Influencing Variables for Rural State Highway Model .....	124
Table 7.3 Rural non-State Highway Model .....	130
Table 7.4 Summary of Influencing Variables for Rural non-State Highway Model .....	131
Table 7.5 Urban Roads Model .....	136
Table 7.6 Summary of Influencing Variables for Urban Roads Model .....	137
Table 7.7 Immediate vs. Influence Ratio Statistics .....	146
Table 11.1 Most Common Movement Types for NZ Fatal Crashes 1996-2005 .....	169
Table 11.2 Pearson Product Moment Correlation Coefficients and p-values ( <i>italicised</i> ) for Rural State Highway Variables, part A .....	170
Table 11.3 Pearson Product Moment Correlation Coefficients and p-values ( <i>italicised</i> ) for Rural State Highway Variables, part B .....	171
Table 11.4 Pearson Product Moment Correlation Coefficients and p-values ( <i>italicised</i> ) for Rural non-State Highway Variables, part A .....	172
Table 11.5 Pearson Product Moment Correlation Coefficients and p-values ( <i>italicised</i> ) for Rural non-State Highway Variables, part B .....	173
Table 11.6 Pearson Product Moment Correlation Coefficients and p-values ( <i>italicised</i> ) for Urban Road Variables, part A .....	174
Table 11.7 Pearson Product Moment Correlation Coefficients and p-values ( <i>italicised</i> ) for Urban Road Variables, part B .....	175



## Glossary

Term	Definition
AADT (or <i>DT</i> )	Average Annual Daily Traffic
Arc	A collection of links between two adjacent nodes.
<i>BD</i>	Bend Density index (number of bends per kilometre of road)
Bendiness	the cumulative variation in horizontal direction along a length of road
Bendiness measure	A way of gauging the degree of bendiness along a stretch or road or for a road network
Binary Logistic Regression	A form of statistical model with two possible outcomes and several influencing variables
<i>CA</i>	Cumulative Angle (turned per kilometre)
CAS	Crash Analysis System (used in New Zealand for recording characteristics of traffic crashes)
Comparison Location	A vertex that does not lie within the influence area of any crash on the road network used.
<i>DR</i>	Detour Ratio (ratio of actual road distance to straight distance between nodes)
External Vertex	A vertex immediately beyond a crash (or comparison) location's influence area.
GIS	Geographical Information System
<i>I</i>	Intersection within 30m dummy variable ( $I = 1$ if the location is within 30m of an intersection, $I = 0$ otherwise)
Immediate Area	A zone of travel distance surrounding a crash (or comparison) location that much smaller than the influence area (taken as 250m for this research's case study)
Influence Area	A zone of travel distance less than or equal to a specified value from a crash (or comparison) location (taken as 1000m for this research's case study)
<i>JD</i>	Junction density (junctions per km)
<i>L</i>	length (km) of a segment
Link	A straight section of digitised road.
<i>Lt</i>	Length of a tangent to a curve
LTNZ	Land Transport New Zealand

<i>MA</i>	Mean Angle (of each bend in the network)
<i>MB</i>	Area of location's meshblock (km <sup>2</sup> )
<i>ME</i>	Mean Elevation above sea level (m)
Meshblock	Unit of analysis for Census information
<i>MR</i>	Mean Rainfall (mm/month)
Node	Location on a digitised road network where more than two links join or a boundary occurs
<i>PE</i>	Number of people residing plus number of people employed within the location's meshblock divided by the meshblock area (people/km <sup>2</sup> )
<i>R</i>	Radius of a horizontal curve (m)
Route	A course of travel between a crash (or comparison) location and one of its external vertices.
<i>SD</i>	Standard Deviation of Angles (degrees)
<i>SI</i>	Straightness Index (proportion of straight lengths in the network)
TLA	Territorial Local Authority
Vertex	Location on a digitised road network where two or more links join
<i>Y</i>	the probability of a crash occurring ( $0 \leq Y \leq 1$ )
$\theta$	deflection angle (degrees) between two links

## 1.0 Introduction

Road related crashes are generally considered to be caused by three major factors: the road environment (the design characteristics of the road itself); the road user (due to decision making and human error); and the vehicle (mechanistic properties). In some cases one of these factors may be seen as the sole cause of a crash, in others it may a combination.

Studies suggest that the road user is to blame from the majority of crashes, with road environment considered the sole contribution in only about 2% of crashes (Grime 1987). This would suggest that traffic engineering, which focuses on attributes of the road environment, might not be the most effective method of reducing crashes. However, as stated by Grime (1987) “road engineering improvements can play a large part in reducing crashes where the road user fails to cope with the road environment” and crashes where the user and the road environment are seen as co-contributors account for about 24% of the total. Thus, engineering measures can be employed to improve the road environment so that road users are less likely to make errors and lower the severity of crashes that still occur.

In a study of traffic crashes in Britain, Haynes *et al.* (2007a) focused on the effects of one contributing road environment factor, the “bendiness” or cumulative horizontal curvature throughout a road network. The study concluded that, contrary to most expectations, bendiness actually had a protective effect on traffic and the regions with a greater degree of bendiness had the fewest traffic crashes when all other significant factors were taken into account.

A similar study was conducted in New Zealand (Haynes *et al.* 2007b) but found very little indication of a relationship between bendiness and crash rate. The only significant trend observed was that, for urban roads, certain bendiness measures indicated bendiness to be protective.

From a traffic engineering point of view, it seems that the aggregation level of these studies was too large but the general technique and measures used may be of use when applied to more localised regions.

This research stemmed from Haynes *et al*'s (2007b) New Zealand study (which for the purposes of this report shall be termed “the motivating study”) in an attempt to further explain the effect of bendiness on traffic crashes in New Zealand. It was based on the hypothesis that a more appropriate method could be developed.

The refined objectives of this research, based on the initial desire to improve the methods of the motivating study and subsequent understanding gained during the course of the research, were to:

- Improve the method of analysis used in the motivating study by providing alternative measures of bendiness, considering the importance of influencing variables and taking into account factors such as flow choices at intersections;
- Determine the relationship between network bendiness and traffic crashes in New Zealand using the revised measures and methods;
- Compare and contrast results with those of the motivating studies study and Haynes *et al*'s (2007a) British study; and
- Determine relevancy of findings to New Zealand practices for road design.

The research method used to achieve these objectives was somewhat organic as the direction taken often depended on results gained from preliminary conclusions. In a sense, the intention of the method was to explore the data rather than follow a rigid pre-determined structure.

The first stage of the research, once motivation was gained from Haynes *et al*'s (2007b) New Zealand study, was to review previous studies that related to the effect of bendiness on traffic crashes, use of GIS in network analysis and associated statistical methods. Next, a more in-depth review of the motivating study and its British counterpart (Haynes *et al.* 2007a) ensued, concentrating specifically on the appropriateness of the methods used and possible changes that could be made. From this, a general plan of the new measures and methods to be tested was developed.

Collation and preparation of data were necessary before any calculations could be performed. This step was also necessary at later stages of the process when needs for different data were recognised.

A preliminary comparison of results using the new methods developed was made with respect to the results of the motivating study. This provided validation to continue the research and a more clearly defined direction of where to proceed. Next, an investigation of possible influencing variables not relating to bendiness was performed and a more thorough application of the method including consideration of these new variables was undertaken. Limitations of the method were assessed and improvements made where possible.

The final stage of the process was to analyse the implications of the research on New Zealand's road design practices and make suggestions about how the developed measures and method could be improved. Suggestions were made for further studies that could enhance traffic engineers' understanding of the effect of road network bendiness on traffic crash occurrence.

## 2.0 Literature Review

This section investigates research on the relationship between bendiness and road-related crashes already undertaken. Sub-section 2.1 outlines general background information with special attention given to those studies interested in cumulative curvature (bendiness) in road networks, design consistency and the use of Geographical Information Systems and statistical methods. Sub-section 2.2 focuses specifically on review of the two previous studies from which this thesis research stemmed directly. An overall summary of the findings from the literature review and the implications for this research is then presented in sub-section 2.3.

### 2.1 General Background Literature

#### 2.1.1 Curves as Contributing Crash Factors

From the characteristics that comprise the road environment, Nicholson (2006) identified seven key aspects that affect crash occurrence:

- Horizontal curvature (generally a function of radius,  $R$ , as illustrated in Figure 2.1 below);
- Vertical curvature;
- Tangent length;
- Grade steepness and length;
- Sight distance;
- Coordination of horizontal and vertical alignments; and
- Overall geometric standard.

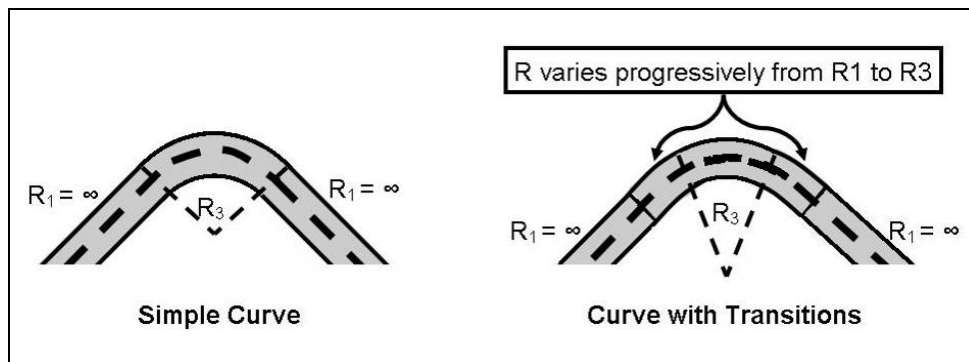


Figure 2.1 Two Types of Horizontal Curvature (adapted from Vis (2007))

While it may be difficult to quantify each of these factors' contributions to crash occurrence, horizontal curvature is noted to be a sole or partial cause of many traffic crashes throughout the world. New Zealand crash data (Ministry of Transport 2007) suggests that approximately half of all fatal crashes in the past ten years occurred in a region of horizontal curvature.

Similar rates have been observed in America; Torbic *et al.* (2004) stated "42,815 people were killed in 38,309 fatal crashes on the U.S. highway system in 2002. Approximately 25 percent of these fatal crashes occurred along horizontal curves". Previously, Lamm *et al.* (1992) had estimated this to be greater than 50%. Analysis of crash types has shown that the large majority of curve-related crashes involved single vehicle, run-off-road type manoeuvres and a much smaller number involved head-on collisions with opposing vehicles (Torbic *et al.* 2004).

Traditionally, determining the effect of horizontal curvature on traffic crashes has been limited to studies of individual bends at discrete locations. The general conclusion of many such studies is that the occurrence of crashes increases with increasing degree of curvature (or decreasing radius of curvature) when all other contributing factors are held constant. However, sections of road with infinite curvature (i.e. straight sections) have been shown to have the same crash rates as medium-sized curves (Gibreel *et al.* 1999).

Grime (1987) attributed this observation to the fact that the level of speed at which a vehicle is likely to skid or lose control is lower for a curve than a straight section, especially when the road surface is wet. Misleading road alignments and poor sight distances associated with geometry were other contributing factors identified. It was also noted that some road alterations that enabled faster speeds on curves but did not involve realignment resulted in increased crash frequency.

Gibreel *et al.* (1999) reasoned that a higher degree of curvature results in a road having greater restriction on driving manoeuvres and hence increases the likelihood of crashes occurring. However, Hauer (2000) debated whether the higher crash frequency observed on sharp curves was due to the degree of curvature or was related to the "point risks" at the entry and exit of the curve. If the latter is true, it would be

only the start and end of the curve, not the length of the curve itself that affects crash occurrence.

Ikeda and Mori (2005) found that the decreasing rate of crashes with increasing radius of curvature occurred only to a point. For radii of curvature greater than about 500m crash rates increased with increasing radius and hence straight roads were among the highest in crash rates. Different crash types were analysed in the study and it was found that rear-end collisions were generally associated with curves of greater radius whereas crashes involving movements such as head-on collisions, vehicle-object collisions and vehicle-pedestrian collisions were generally associated with curves of lesser radius.

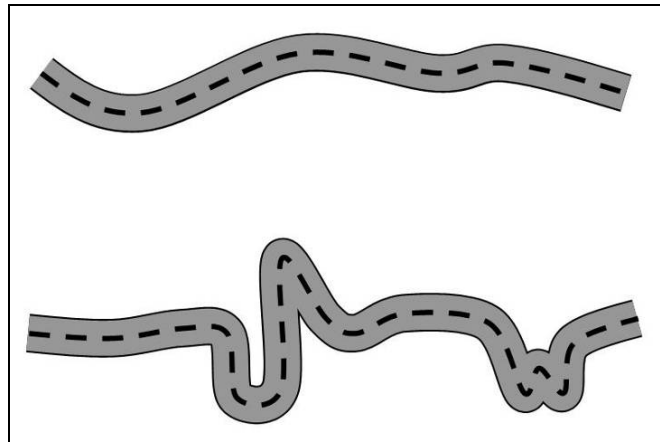
The limitation of studies that focused on individual curves to obtain relationships between crash frequency and degree of curvature is that they did not take into account the wider context of the road environment surrounding the individual curves. Hauer (2000) classified the factors that contribute to road safety as either *internal* or *external*. Internal features are those characteristics inherent to an individual road section, for example its degree of curvature or superelevation and were the focus of the studies that examined individual curves. External features are those characteristics that influence driver perception and approach speed, for example the density of curves upstream, length of connecting tangents and available sight distance. This classification brings a distinction between the properties of an individual curve and its location relative to the road network in which it is situated.

Many other researchers have noted the prominence of the context in which a curve is located. Over forty years ago English traffic engineers noted that roads with long straights and few curves generally had higher crash rates than similar roads with many curves (Road Research Laboratory 1965 as cited by Nicholson, 2006). This was supported by Wilson (1968) who identified a danger in having a single sharp curve after a long tangent. Milton and Mannering (1996) stated that the geometry of horizontal curves is not dangerous but that the placement of curves after long straight stretches of roads will increase crashes. From their research, Noland and Oh (2004) also suggested that roads with many curves may not necessarily be less safe than straighter roads.



### 2.1.2 Bendiness

This gives rise to the notion of “curviness” or “bendiness”, which is traditionally known as the cumulative variation in horizontal direction along a length of road (McLean 1989). Put more simply, measures of bendiness gauge the frequency and sizes of several curves as opposed to examining curves individually. This allows comparison of different road sections, for example in Figure 2.2 below, where the two different roads might both be subjectively termed “bendy” but a more qualitative evaluation would be needed to properly distinguish between the two. Many different measures of bendiness that gauge the proportions and sizes of curves and tangents (straight sections of road) that exist along a stretch of road or for a whole road network have been defined. However researchers have not agreed on any one definition of bendiness that is more appropriate than others.



**Figure 2.2 Plan View of Two "Bendy" Roads**

Among five measures used by Lamm *et al.* (1986) in an evaluation of multiple horizontal and vertical elements for a length of highway was the measure of *average curvature*, defined as “the sum of central angles of horizontal curves in a specific highway section divided by the length of this section”. Another measure used was *length ratio*, defined as “the sum of horizontal and vertical curve lengths in a specific highway section divided by the length of this section” (Lamm *et al.* 1986).

$$AC = \frac{\sum \gamma}{\sum dr}$$

**Equation 1 Average Curvature (from Lamm *et al.* (1986))**

$$LR = \frac{\sum dc}{\sum dr}$$

**Equation 2 Length Ratio (from Lamm *et al.* (1986))**

where:  $AC$  = the average degree of curvature

$LR$  = the length ratio

$\gamma$  = the angular size (degrees) of an individual curve

$dr$  = the total distance (km) along a route

$dc$  = the distance (km) along an individual curve

Shankar *et al.* (1996) analysed the effect of intelligent transportation systems on crash severity by considering many indicators, two of which were concerned with measures of bendiness, although this was not explicitly stated in their analysis. They used the percentage of horizontal curve length per kilometre of roadway, similar to the length ratio measure. It was found that as this percentage increased so too did the likelihood of an injury. The other bendiness measure used was simply the number of horizontal curves per kilometre of roadway. This too was found to be directly related to injury likelihood. This second observation gives merit to Hauer's (2000) postulation that it may be the presence of entry and exit points of a curve, not the curve length, that decreases safety.

Bjorketun (2005) also used the traditional bendiness measure of average degree of curvature (the sum of successive absolute changes per kilometre). It was found that, for high-speed rural locations in Sweden, hilliness (the vertical curvature counterpart of bendiness) was more significant in causing crashes.

Castro *et al.* (2005) used six different horizontal alignment indices to gauge the effect of bendiness on crash rates. The first index, *curvature change rate (CCR)*, was the traditional bendiness measure, the sum of deflection angles divided by the total road length.

$$CCR = \frac{\sum \Delta\theta}{L}$$

**Equation 3 Curvature Change Ratio (from Castro *et al.* (2005))**

where:  $\theta$  = the deflection angle (degrees)

$L$  = the length (km) of a segment

*Degree of curvature (DC)* used a fixed horizontal curve length of 100m divided by the radius of each individual curve, segments were then evaluated by summing the individual degrees of curvature and dividing by the total segment length.

$$DC = \frac{\sum DC_i}{L}$$

**Equation 4 Degree of Curvature (from Castro *et al.* (2005))**

where:  $DC_i$  = the degree of curvature of curve  $i$

It should be noted that the term “degree of curvature” is also commonly used to express bendiness as deflection per 100 feet or 100 metres.

*Curve length: Roadway length (CR)* was the total length of curved sections in a segment divided by the total overall length of the segment.

$$CR = \frac{\sum CL_i}{L}$$

**Equation 5 Curve Length: Roadway Length (from Castro *et al.* (2005))**

where:  $CL_i$  = the length of curve  $i$

*Average radius ( $\bar{R}$ )* gauged segments according to the mean radius of curvature when considering all curves in the segment.

$$\bar{R} = \frac{\sum R_i}{Nc}$$

**Equation 6 Average Radius (from Castro *et al.* (2005))**

where:  $R_i$  = the radius of curve  $i$

$Nc$  = the number of horizontal curves within a segment

*Average tangent* ( $\bar{T}$ ) computed the average length throughout a segment of straight sections leading up to horizontal curves.

$$\bar{T} = \frac{\sum TL_t}{Nr}$$

**Equation 7 Average Tangent (from Castro *et al.* (2005))**

where:  $TL_t$  = the length of tangent  $t$

$Nr$  = the number of tangents within a segment

*Maximum radius/ minimum radius (MR)* the final index used by Castro *et al.* (2005) examined the ratio between the greatest and lowest radii of curves for a segment. This was considered to be a good indicator of homogeneity but did not gauge bendiness well, as the same ratio could be contained for segments of very different curvature.

$$MR = \frac{R_{\max}}{R_{\min}}$$

**Equation 8 Maximum Radius/ Minimum Radius (from Castro *et al.* (2005))**

where:  $R_{\max}$  = the maximum radius of curvature within a segment

$R_{\min}$  = the minimum radius of curvature within a segment

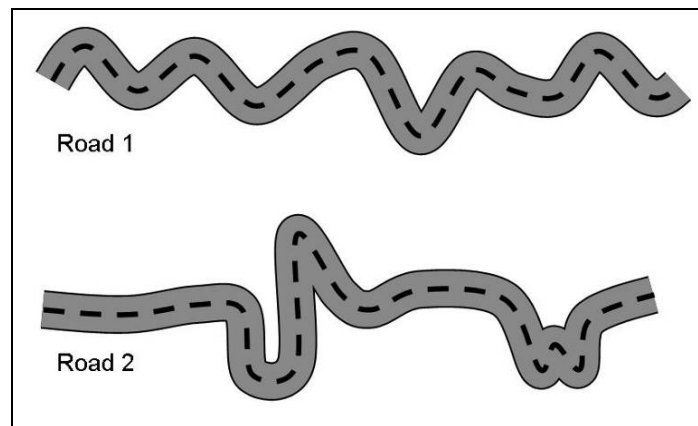
Castro *et al.* (2005) found that the curvature change rate gave the best correlation between bendiness and crash rates with  $R^2 = 0.60$ , however, this could be improved to  $R^2 = 0.66$  when a composite index including both horizontal and vertical curvature change rates was used.

Much information is available on the design theory for individual curves, where forces due to superelevation, degree of curvature, friction and vehicle weight are combined to give an indication of the safe travelling speed around a curve (for further detail see Nicholson and Saleh (2004) or Austroads (2003)). However, no literature

could be found where this type of mechanical theory had been extended to design for bendiness rather than just isolated curves.

### 2.1.3 Design Consistency

The concept of design consistency is another approach used to illustrate the impact of variability of successive road elements on crash occurrence. Rather than purely considering the number of horizontal curves per given length of road, as is the case for some bendiness measures, it is important to consider the placement of each curve relative to others and the variations in the speeds at which vehicles can travel along each element. For example, Figure 2.3 shows two roads, Road 1 would be classified as more bendy than Road 2, according to the definitions of bendiness discussed previously. Also, Road 1 is obviously designed more consistently than Road 2.



**Figure 2.3 Plan View of Two Sample Roads to Illustrate Design Consistency**

Observations that the location of horizontal curves relative to other road elements has a great effect on crash occurrence illustrate the importance of design consistency. To achieve design consistency is the aim of most geometric design guides worldwide as it “ensures that successive geometric elements act in a coordinated way, so that they produce harmonized driver performance without surprising events”(Gibreel *et al.* 1999).

Hauer (2000) attributed the high crash frequencies of sharp curves that follow long straight tangents to driver behaviour and the road’s unexpectedness. Hence

successive elements that are substantially different are likely to cause driver confusion.

Elvik and Vaa's (2004) meta-analysis of 12 international studies on the effects of horizontal curve treatments on crashes supported Hauer's observations. They stated, "While driving on country roads, the driver forms expectations of the trajectory of the road on the basis of the road alignment. When the road is, in the main, straight, the driver does not expect sudden sharp curves to occur. When the road has numerous curves, on the other hand, the driver will expect there to be further curves on the road ahead."

Polus and Mattar-Habib (2004) confirmed the importance of alignment consistency but thought that most designers did not take this into account. They believed that having less variability between elements results in a more even speed distribution for each driver as they travel along the road and this in turn should result in lower driver "strain" and improved safety. However, Nicholson (1998) noted that the goal of having a small variation in the margin of safety of successive curves conflicts with another important goal of ensuring a substantial margin of safety at each curve. Alterations to make individual curves more similar to surrounding curves (i.e. increasing the mean margin of safety) may often lead to an increase in the variation of the margin of safety of the curves. Thus achieving design consistency is a difficult process which requires compromise between objectives.

Approaches to gauging consistency generally focus on speed measurements. Horizontal curves, like all road elements, have a design speed, which is "a selected speed used to determine the various geometric design features of the roadway", (AASHTO 2001). Design speeds should be at least equal to the 85<sup>th</sup> percentile of the distribution of operating speed, the "the speed of cars at a time when traffic volumes are low and will allow a free choice of speed within the road alignment" (Austroads 2003). Lamm *et al.* (1986) suggested that variation in operating speeds of successive elements could be used to approximate actual design speeds and provide a simple gauge for evaluating design consistency. Gibreel *et al.* (1999) supported this notion by saying that "for successive geometric elements, design consistency is evaluated based on the operating speed on these elements." Park and Saccomanno (2006) also

recognised this by noting that “in principle, safe highways are those highways that maintain consistency of vehicle speeds between the upstream and the downstream elements (e.g. a tangent/curved section).”

Jackett (1992), in an analysis of New Zealand’s advisory speed sign policy, showed that the probability of a crash occurring on a curve increased with the percentage difference between the curve’s approach speed (actual speed taken by drivers leading up to the curve) and advisory speed (the speed suggested but not legally expected as a guide for comfortable travel around the curve). In an evaluation of New Zealand crash analysis procedures, Koorey and Tate (1997) confirmed this finding, showing that crash rates increased by 2.5% with every 1km/h difference between curve and approach speeds.

Park and Saccomanno (2006) debated the appropriateness of the use of the change in 85<sup>th</sup> percentile speeds to approximate design consistency. Their research, based on that of McFadden and Elefteriadou (2000) proposed a new measure of design consistency, the “85MSR”, which reflects the maximum speed reduction between successive elements as experienced by the same vehicle or driver.

Whereas the 85th percentile model assumes that vehicles travelling over road sections maintain the same ranking in terms of their speed, the 85 MSR model assumed that the speed profiles of individual vehicles on successive elements lie somewhere between a perfect positive correlation (where the fastest vehicle on a tangent is also the fastest on the following curve) and a perfect negative correlation (where the fastest vehicle on a tangent is slowest on the following curve). Analysis of a data set containing 18 tangent-curve pairs from two rural highways found that the speed differentials calculated using the 85MSR approach were on average 1.6 times greater than the results of the standard change in 85<sup>th</sup> percentile approach.

It was concluded that the 85MSR approach was a more realistic model of how drivers actually chose their speeds and was more conservative and therefore more suited to determining potential crash risk. It was also concluded that more cost-effective alternatives to changing geometric alignment in order to reduce individuals’ speed differentials and hence improve safety might exist. (Park and Saccomanno 2006)

Polus and Mattar-Habib (2004) developed two new measures of consistency in their research on two-lane rural highway safety. The first measure, the relative area, measured the extent to which speeds on individual elements differed from the mean speed over the whole road section, i.e. the spread of speeds, by calculating the area bounded by the speed profile and average operating speed on a speed versus distance graph. The second measure, also concerning the spread of speeds, was the standard deviation of all speeds along the road section. These methods were considered advantageous as they considered whole highway sections rather than just pairs of elements as with the 85MSR measure. It was observed that study sections with higher consistency values had lower crash rates.

Koorey (2005) presents a collection of equations that can be used to determine design speeds and hence 85<sup>th</sup> percentile operating speeds and speed profiles for sections of road when the road geometry data is known. This enables an alternative to on-site surveying, as was performed in many of the design consistency studies presented here, as only geometry data is required. By using this method of obtaining design speeds it should be possible to determine speed consistency measures for much larger areas, as long as the geometry data provided was consistent and accurate. This method was used to contrast the “local speeds” of curves with their surrounding speed environments to identify potentially hazardous locations where the two measures were poorly matched.

Echaveguren *et al.* (2005) suggested that consistency measures that calculated margins of safety based on operating and design speeds are unreliable as design speed calculations do not properly account for the balance between friction demand and provision. They proposed a method of measuring the reliability of horizontal curves based on probability distributions of friction supply. When applying their reliability method to observed data in a case study of five horizontal curves and then using a sensitivity analysis of different variables it was found that curve radius, skid resistance and macrotexture had the biggest effect on reliability.

Norwegian researchers use a computer program first developed in 1984 called the URF program to identify the risk of driving off the road due to unexpected curves



(Elvik and Vaa 2004). The URF risk value depends on a curve's degree of "unexpectedness", road width, and road gradient. The unexpectedness in turn depends on the curve's difference in driving speed, radius, and superelevation, all compared with average values for a larger road section. It was found that the URF risk for a road section with 0.5 or less curves per kilometre was three times greater than a road section with 0.75 or more curves per kilometre.

Castro *et al.* (2005) proposed a consistency rating using a qualitative scale of either good, fair or poor based on values of curvature change rate, vertical curvature change rate or a composition index of the two variables. However, they suggested that these characteristics did not evaluate consistency as well as measures that consider operating speeds.

Whatever method used in estimating design consistency is used, approaches aimed at increasing design consistency must be carefully considered. Elements designed to reduce the effects of a sudden change in design speed, for example spiral transition curves between straight tangents and circular horizontal curves may actually decrease safety. Nicholson (2006) points out that transition curves may affect drivers' perceptions of the true road geometry and lead to the use of unsafe speeds. By increasing the radii of several successive curves to allow greater speeds, the overall speed travelled along the road may become unsafe, especially if additional curves downstream have not been treated in a similar way (AASHTO 1990).

Figuerola and Tarko (2005) defined design consistency as "the conformance of the highway geometry with driver expectations." While they acknowledged that the assessment of operating speeds versus design speeds or the variation in speeds between successive elements are ways of gauging design consistency, their definition brings another aspect that previous definitions did not, that is the expectations that drivers have of the road environment based on their own experiences and perceptions.

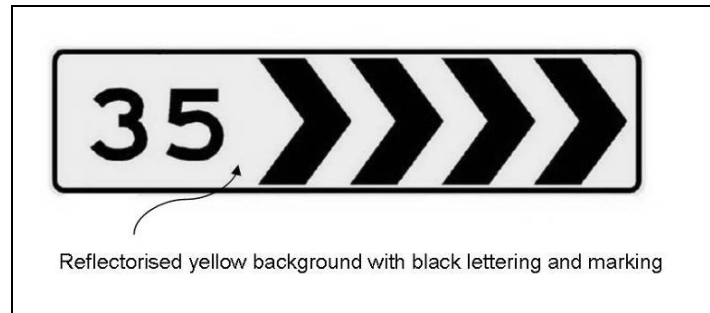
Other geometric features may contribute to drivers' perceptions of horizontal curves. Hassan and Easa (2003) used on-site speed observations and surveys involving 3D animations to determine the effects of combining horizontal and vertical curves. They found that horizontal curves combined with vertical crest curves were perceived as

sharper than the same sized horizontal curve combined with a vertical sag curve, regardless of other factors such as superelevation or turning direction. Drivers tended to decelerate for crest/horizontal curve combinations but would accelerate at the beginning of a sag/horizontal curve combination, regardless of other geometric variables.

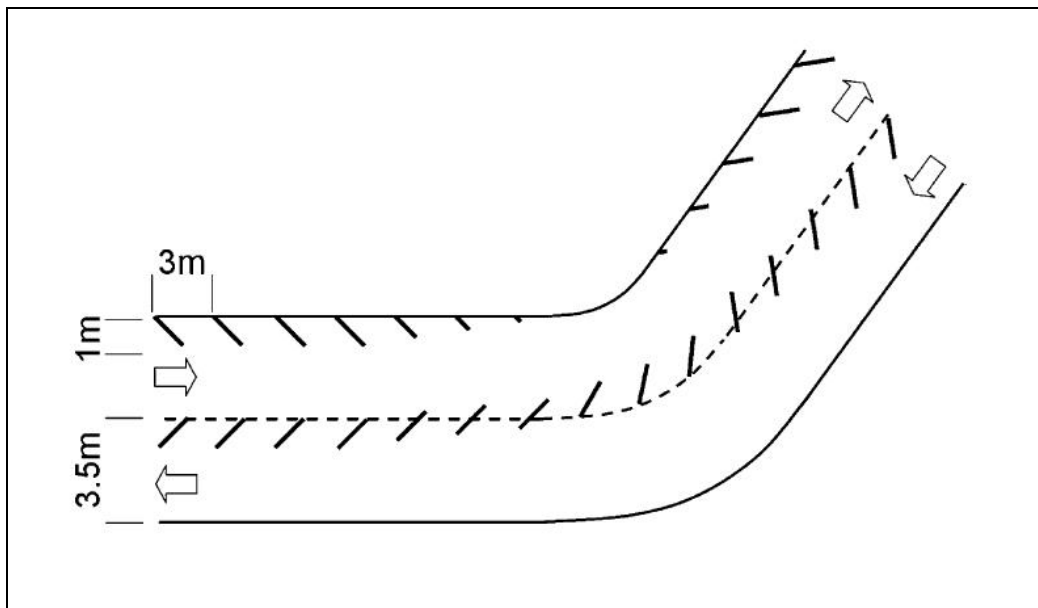
#### 2.1.4 Driver Awareness of the Road Environment

As well as geometric features, treatments performed on horizontal curve sites can affect drivers' perceptions of the risks involved when travelling around curves. Vest and Stamatiadis (2005) studied the effects of warning signs (suggested speed limits with additional flags or flashing lights and repeated along the road) and modified pavement markings (including extra delineator posts and transverse pavement lines) used to give drivers information regarding approaching horizontal curves in areas of design inconsistency. Their literature review suggested that both signs and pavement markings would reduce operating speeds and hence increase safety. It was found that all treatments trialled reduced variability of operating speeds but some treatments increased the operating speeds at individual sites. The overall conclusion was that warning signs and pavement markings do affect drivers' choice of speeds. This is important as it shows that there are many factors that can contribute to traffic safety in bendy areas other than the properties of geometric features.

Charlton (2007) in a study of New Zealand drivers' reactions to advance warning, delineation and road marking treatments to horizontal curves found that chevron signs (Figure 2.4) and herringbone markings (Figure 2.5) were the most effective methods of reducing drivers' speeds and improving their positioning on the road. It concluded these treatments were the most effective as they gave perceptual cues and therefore increased driver attention, decreased the likelihood of misperceptions and guided drivers to better lane positioning.



**Figure 2.4 Chevron Curve Warning Sign (from Transit New Zealand *et al.* (2007))**



**Figure 2.5 Herringbone Pavement Marking (from Charlton (2007))**

Elvik (2006) proposed four rules of accident causation, two of these are of specific interest to this study. *The law of complexity* implies that a greater number of potentially relevant items of information requiring a driver's attention increases the crash risk. Also, *the law of cognitive capacity* implies that challenges to a driver's mental capacity will be more influential than challenges to a driver's physical ability. Thus the road environment presents difficulties to the driver. If these difficulties result in impairment of mental functioning then a crash is more likely to occur.

Similarly, Mahalel and Szternfeld (1986) showed that crashes occur when the demand placed on a driver by the nature of the road environment is greater than the driver's level of awareness (or performance). Therefore, changes in the demand of the road

environments that occur quickly or without sufficient warning are likely to lead to crashes. When designing the road environment it is important to consider the ability of drivers to cope with the demand it places upon them.

Messer *et al.* (1981) gave four characteristics of a geometric feature that affect the mental workload it imposes on a driver:

- The criticality of the feature (function of feature type, frequency, complexity and crash potential);
- The sight distance to the feature;
- The similarity between the feature and the preceding feature; and
- The level of driver familiarity with the feature.

#### 2.1.5 Demand of the Road Environment on Drivers

While engineering design aims to reduce the level of demand placed upon the user by the road environment in order to avoid crash occurrence (Grime 1987) there is much evidence to suggest that reducing the level of demand too much can have negative implications on road user safety. Mahalel and Szternfeld (1986) noted that engineering treatments that make driving tasks easier may result in a reduction of attentiveness to a level that is unsafe and will result in increased crash frequencies. They also warned that, as an increase in speed results in a decrease in safety, treatments that improve safety of elements but increase driver speeds might also increase crash frequencies.

An inability to cope with the task of driving may simply be due to driver inattentiveness. In Stutts *et al.* (2005), driver inattention was identified as a contributing factor to 25-30 percent of crashes on American highways. Inattentiveness was contributed to the driver being either fatigued or distracted through visual, auditory, physical or cognitive means.

Ikeda and Mori (2005) attributed the increase in rear-end crashes with increase in radius of curvature to driver inattentiveness and misperceptions of safety. They suggested that drivers should be stimulated by other measures in areas where the road alignment does not provide much challenge to the driving task.

Gibreel *et al.* (1999) explained the observed phenomenon of crash rates for sections of road with zero curvature having higher crash frequencies than sections with moderate curvature by saying that simple tasks decrease driver arousal. They highlighted the difficulty of trying to balance the goals of simplifying the demand of the road environment and maintaining an adequate level of drivers' perception of difficulty and performance.

#### 2.1.6 Use of GIS Techniques

The previously explained observations by Mahalel and Szternfeld (1986) and Gibreel *et al.* (1999) are contrary to those of Cairney (2005) who observed that crash rate was proportional to the demand of the combinations of geometric features along a road section which was gauged by the advisory speeds for road sections. Cairney's study on the relationships between geometric alignment and crashes made use of GIS to combine data from different sources spatially. The data sources used were: road geometry information from the GIPSI-TRAC road geometry measurement system; cross-section data from visual inspections; traffic flow data from road authority records; and crash data from road authority databases.

In terms of curvature, the study concluded that extreme values (i.e. very sharp curves) increased crash rates substantially. GIS was seen as a critical component of the study as it allowed links between geographical location and contributing characteristics and analysis of crashes in relation to spatially varying features such as traffic flow levels or geometric properties. (Cairney 2005)

Noland and Quddus (2004) also used GIS in their spatially disaggregate analysis of road casualties in England. This enabled them to identify variables specific to certain areas when examining the crash characteristics for each of England's 8414 census wards. Possible further uses of the GIS such as spatial autocorrelation techniques (that measure the degree of clustering between events), the use of time series data and further disaggregating of the crash database used to examine more potential contributing variables were identified. Although the focus of this study was not related to horizontal curvature it provides a good example of GIS use in a transportation-related study.

One American study used data collected by instrumented vehicles to determine the characteristics of road sections (Drakopoulos and Ornek 2000). A GIS program classified road sections as either tangent, horizontal curve to the left, or horizontal curve to the right and determined the degree and length of curves based on information on distance travelled, gradient, transverse slope and compass bearing collected by instrumented vehicles. The biggest problems faced in developing the program came in determining the point of tangency (PT) or point of curvature (PC) for elements; errors made in this would lead to poor estimates of curve parameters. It was postulated that the developed software could be used to identify hazardous geometric features (example given – long tangents followed by curves with large degree of curvature or successive curves with different radii) from crash data and then screen network for similar sites that have not yet experienced high crash rates but should be treated.

Many previous studies that used GIS but had nothing to do with road safety can still be applied to this research. One such example is that of Steven *et al.* (2004) who used GIS to identify the common habitat characteristics of locations where Canada's endangered fish, the Topeka Shiner, had been observed. These characteristics were then used in a multivariate logistic regression to identify other potential sites of similar habitat quality that could be prioritised for conservation interventions and protection. This is a good example of a study that, given known locations of events (observations of Topeka Shiner schools) in a network (the natural water system) identified the characteristics of the locations and searched the network for other locations with similar characteristics.

Another similar study by Stephens (2004) studied the habitats of successful duck nesting sites in the Missouri Coteau Region of North Dakota. Characteristics of these sites were modelled using generalised linear regression and this model applied to a larger region to guide conservation programmes.

Many other studies have used existing models to identify locations of high risk. For example, given factors favouring the ignition of a forest fire (such as combustion parameters, inflammable material, slope, proximity to the road network, and urban

areas and distance from water sources) Petrakis (2005) used GIS to identify other likely fire hazard areas in one of Greece's national parks. Similarly, Wallis (2005) identified areas throughout Mississippi that had high potential for a West Nile virus outbreak given mosquito breeding habitat characteristics.

Wood (2003) used ArcView's Network Analyst to determine, for all the staff working at Vodafone's headquarters in Newbury, England, the most likely route travelled from their home to work. This information was then used in the GIS to determine the volume of CO<sub>2</sub> emissions attributable to and amount of road infrastructure used by Vodafone's staff in order to compute the company's ecological footprint and produce a sustainable business travel plan. The use of GIS was seen as integral to the study as, without it, estimation of the routes taken by staff members would have been an extremely time intensive task.

#### 2.1.7 Level of Aggregation

Montello *et al.* (2001) defined three levels of scale used in geography: *cartographic scale*, which is the ratio of a feature's represented size on a map to its actual physical size; *analysis scale*, which is the size of units at which a problem is analysed; and *phenomenon scale*, which is the size at which structures or processes physically exist. It was stated that "in order to observe and study a phenomenon most accurately, the scale of analysis must match the actual scale of the phenomenon." However, this is not always possible as, in order to collect real-world data and represent it digitally, the data must be aggregated (or "generalised") to a certain extent.

When aggregating phenomena and analysing differences between areas, the definition of the boundaries that make up these areas is also very important, because different groupings of phenomena can result in very different analysis outcomes; this is known as the "modifiable areal unit problem" (MAUP). Manley *et al.* (2006) explained that the MAUP involves two distinct issues; definition of boundaries and scale effects.

Harris and Johnston (2003) when studying the best method of aiding socially marginalised populations in Britain and Wales asked the question "how should the neighbourhood be defined and at which geographical scale?" They tested the use of

different spatial boundaries to define areas and concluded that a fine-scale geography is best suited to a fine-scale policy.

Manley *et al.* (2006) investigated the effects of the MAUP on studies of two different British census variables, proportion of the population that is female (assumed to be relatively evenly spaced throughout the study area) and the percentage of the population renting local authority-owned housing (assumed to be varied throughout the study area). They concluded that the British census areas were defined according to a particular process (number of people within the areas) and was not suited to the processes of other variables, for example house rental levels. They identified that compromise must always be made as it would never be possible to define areas that capture all processes of different variables.

Grubestic (2007) made a similar conclusion for the zip code system used in the United States of America. The zip codes were established by the United States Postal Service (USPS) as a means of efficiently directing mail but have since been adopted by market researchers as a means of grouping consumers. It was found that as zip code areas are based solely on geography, the demographic characteristics of people within them were generally non-homogenous. Another problem noted was that the USPS was constantly updating and modifying the zip code classification which caused error for long-term comparative studies.

#### 2.1.8 Statistical Evaluation of Results

The type of statistical evaluation used in crash studies can have a large effect on the conclusions drawn from the data (Elvik and Vaa 2004). Abdel-Aty and Radwan (2000) state that three methods of predicting crash occurrences with respect to contributing variables have been used by previous researchers. These methods are multiple linear regression, Poisson regression and negative-binomial regression.

For a case where a several explanatory (or “predictor”) variables can be linearly related to a dependant variable, multiple linear regression models apply weighting factors to these individual linear relationships and sum the terms. For example:



$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

**Equation 9 Standard form of Multiple Linear Regression (from Nicholson (2006))**

where:  $Y$  = Dependant variable (e.g. number of crashes)

$X_1, X_2$  = explanatory variables (e.g. bendiness, traffic flow)

$\beta_0, \beta_1, \beta_2$  = constant parameters

Alternatively, Poisson regression (where the variance is equal to the mean) and negative binomial regression (an extension of the Poisson case where the variance and mean do not have to be equal), use multiplicative models, for example:

$$Y = \beta_0 X_1^{\beta_1} X_2^{\beta_2}$$

**Equation 10 Standard form of Poisson or Negative Binomial Regression Models  
(from Nicholson (2006))**

Miaou and Lum (1993) showed that multiple linear regression is not a suitable method as it relies on the normal distribution which is symmetrical whereas crash occurrence is infrequent and therefore positively skewed with the majority of road sections having zero crashes during a given observation period. Multiple linear regression also requires sample sizes too large for crash studies.

Abdel-Aty and Radwan (2000) concluded that in cases of over dispersion, where the variance of the data exceeds the mean value, and cases where independence between data sets was not guaranteed, negative binomial regression was considered to be more appropriate than Poisson regression.

Elvik and Vaa (2004) identified that use of Poisson or negative binomial models is the most appropriate method of avoiding over-fitting of data, where effects of the random nature of crash occurrence are sometimes ignored. Hauer (2004) agreed that Poisson and negative binomial models are appropriate but only in situations where traits of entities, including traffic, remain relatively constant.

Maher and Summersgill (1996) described how problems such as low mean values, over dispersion, time effects, random errors and combining site observations generally

associated with generalised linear models (GLMs) can be combated by making changes to the basic GLM technique to develop Poisson and negative-binomial models. In situations where the mean and variance are of similar order, the Poisson distribution was shown to be the most suitable, but when the variance was greater than the mean the negative binomial distribution is most suitable.

## **2.2 Motivating Studies**

Two main previous studies have provided motivation for this research. The following sub-sections describe them in detail.

### 2.2.1 District variations in road curvature in England and Wales and their association with road traffic crashes (Haynes *et al.* 2007a)

In 2005 a comprehensive study of the effect of road bendiness on crash rates for all 403 local authority districts in England and Wales was undertaken (Haynes *et al.* 2007a). The study, hereafter referred to as “Haynes *et al.*’s British study”, was the initial basis for this research.

To represent digitally Britain’s physical roads, data were provided as a series of straight sections termed “links.” Curves were represented by “vertices” where two or more straight sections of road joined. Locations where more than two links joined or a boundary occurred were termed “nodes”. The collections of individual straight sections of road and their vertices occurring between two adjacent nodes were termed “arcs”. These component types are illustrated below in Figure 2.6:

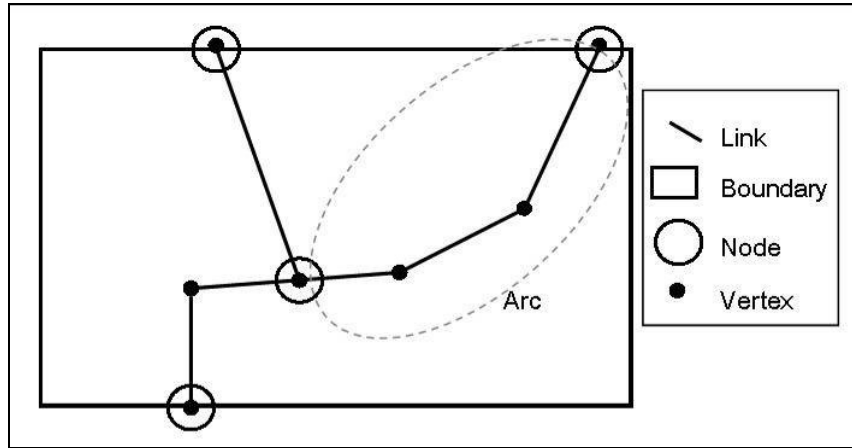


Figure 2.6 Components of Road Network Data (adapted from Haynes *et al.* (2007a))

Haynes *et al.* developed GIS and Fortran programmes to determine, for each district, the bendiness according to five different indicator definitions. These are given with corresponding equations that reference Figure 2.7 where  $a$ ,  $b$ ,  $c$ ,  $d$ ,  $e$  and  $f$  are road links between vertices,  $p$  and  $q$  are direct distances between nodes and  $u$ ,  $v$  and  $w$  are angles between links.

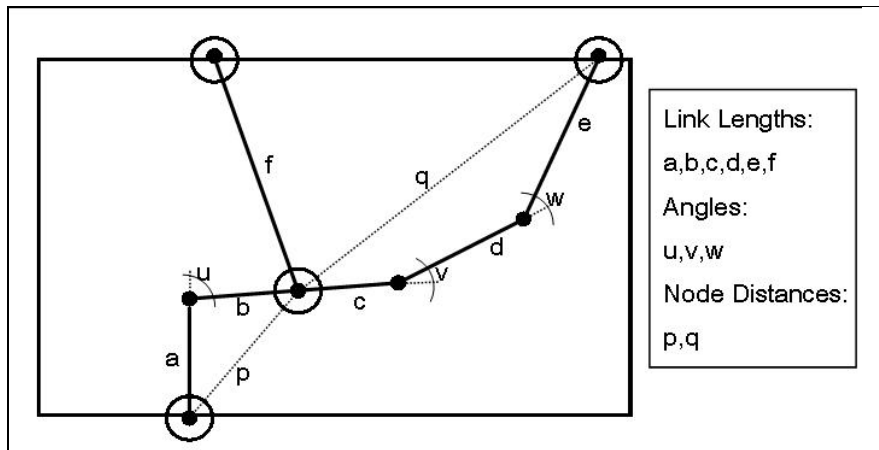


Figure 2.7 Bendiness Diagram (adapted from Haynes *et al.* (2007a))

The bendiness measures used by Haynes *et al.* were:

- Bend density ( $BD$ ) – defined as the number of bends per kilometre of road. This does not include the bends at intersections, i.e. includes only vertices, not nodes, in its analysis. Bend density was calculated with Equation 9:

$$BD = \frac{N_v - N_n}{a + b + c + d + e + f}$$

**Equation 11 Bend Density (from Haynes *et al.* (2007a))**

where:  $N_v$  = number of vertices within the study region

$N_n$  = number of nodes within the study region

- Detour ratio ( $DR$ ) – defined as the ratio of actual road distance to straight distance between nodes, the same as Lamm *et al.*'s (1986) length ratio measure. This was computed for the network as a whole. Due to difficulties associated with approximating curves as a series of straight lines the Haynes studies did not include arcs shorter than 200m in the analysis. By definition the detour ratio must be greater than or equal to one. The detour ratio was calculated according to Equation 10:

$$DR = \frac{a + b + c + d + e + f}{p + q + f}$$

**Equation 12 Detour Ratio (from Haynes *et al.* (2007a))**

- Straightness index ( $SI$ ) – defined as the proportion of straight lengths in the network. This used detour ratios for individual arcs to determine which were straight (i.e. had a detour ratio of 1). Once identified, the length of straight sections was compared to the total length of road in each network. By definition the straightness index must be less than or equal to one and the difference should be equal to the value of the curve length: roadway length ratio defined by Castro *et al.* (2005). The straightness index was calculated according to Equation 11:

$$SI = \frac{f}{a + b + c + d + e + f}$$

**Equation 13 Straightness Index (from Haynes *et al.* (2007a))**

- Cumulative angle ( $CA$ ) – defined as the cumulative angle turned per kilometre. This is the measure used in the study that is closest to the

traditional definition of bendiness and is the same as Lamm *et al*'s (1986) average curvature and Castro *et al*'s (2005) curvature change rate measures. Angles between successive links were computed using compass bearings in a specially designed Fortran programme according to Equation 12:

$$CA = \frac{u + v + w}{a + b + c + d + e + f}$$

**Equation 14 Cumulative Angle (from Haynes *et al.* (2007a))**

- Mean angle (*MA*) – defined as the mean angle of each bend in the network. This was calculated by dividing the sum of all angles by the number of angles in the network according to Equation 13:

$$MA = \frac{u + v + w}{Na}$$

**Equation 15 Mean Angle (from Haynes *et al.* (2007a))**

where:  $Na$  = number of angles between links in the study region.

The crash data studied by Haynes *et al.* (2007a) was obtained for a five-year period from 1995 to 1999 from Police records. Three levels of severity were distinguished: fatal, serious or slight, according to the British crash reporting protocol. Road types, in terms of a hierarchy of “major roads”, “B roads” or “minor roads”, were also identified for each crash location and assessed separately in the analyses.

Possible contributing variables, additional to road bendiness, were identified. Some of these variables aimed to represent exposure to crash risk, such as percentage of population at risk (according to national statistics of ages and gender most likely to be involved in crashes), car ownership, road length and average annual daily traffic flows. Most of the variables were aggregated at a district level but road length and traffic flows were classified for each district according to what type of road (major, B or minor) they occurred on.

Possible variables were tested for significance using Poisson models. The data were found to be overly-dispersed so a negative binomial analysis was used instead. The significant influencing variables were found to be:

- Total length of road in district;
- Annual average daily vehicle movements on all roads;
- District population at risk;
- Number of cars per capita;
- Proportion of road length passing through built up urban land;
- Proportion of road length that is minor roads; and
- Material deprivation index (explanation follows).

Townsend's Material Deprivation Score (Townsend *et al.* 1988), which is commonly used in Britain, examines four factors: unemployment, overcrowding, lack of owner occupied accommodation and lack of car ownership (Local Government Data Unit 2004). The last factor, however, was excluded from the material deprivation index used in Haynes *et al.*'s (Haynes *et al.* 2007a) study as it was accounted for elsewhere.

These variables were held constant through the use of partial regression coefficients in order to examine the effect of each of the five bendiness measures on each of the three crash types. Similarly, crash occurrence with change in bendiness was calculated for each individual road type.

It was determined that bendiness characteristics had distinctive spatial nature; when comparing a bendiness measure over all the different districts large areas of similar values generally appeared.

The bend density indicator appeared to vary differently to the other four indicators, which all had positive associations. A trend for districts that had a high number of bends per kilometre to have bends not as sharp on average was identified.

The cumulative angle was identified as being well correlated to fatal and serious crashes while the detour ratio was identified as being well correlated to serious and

slight crashes. Thus the cumulative angle was chosen as the most appropriate indicator and used in further analysis in order to develop full models with one bendiness measure. It was noted that this indicator did not include all the properties of bendiness that may have been influencing.

Using the cumulative angle model, with all other variables held constant, it was found that an angle increase of one degree per kilometre decreased fatal crashes by 0.57%, serious crashes by 0.71%, and slight crashes by 0.51%. It followed that districts with straight roads were found to have more crashes than those with curved roads.

It was concluded that bendiness was not hazardous but protective on a large scale, although individual bends were still more hazardous than individual straight sections. It was hypothesised that this was due to a combination of decreased speeds on curves, increased driver vigilance and a discouragement of risk-taking behaviour.

Further research aspects, such as determining at which scale bendiness converts from being hazardous to protective and the types of crashes associated with levels of bendiness, were suggested.

#### 2.2.2 Influence of Road Curvature on Fatal Crashes in New Zealand (Haynes *et al.* 2007b)

Another study assessing the effects of road network bendiness on traffic crashes was conducted in New Zealand in 2006 under the guidance of Robin Haynes, a member of the British study team, using the same methodology as Haynes *et al.*'s British study. The New Zealand study, hereafter referred to as "the motivating study" formed the scope of this research.

The motivating study used the same bendiness indicators as for the British study. The crashes selection of crashes to be considered was somewhat different to those of the British study; a ten-year observation period was used, from 1996 to 2005 in order to gain a large enough sample size and only fatal crashes were assessed as it has been identified that New Zealand has poor reporting rates for lesser severity crashes in some regions. Roads were classified as being either urban, rural state highway or rural other and analysis was performed for each separate road type as well as for all

three together. Whereas the regions used for comparison in the British study were based on the 403 local authority districts the New Zealand study data were aggregated according to New Zealand's 73 territorial local authority (TLA) regions.

The other influencing variables trialled were concerned with exposure (a variety of different measures reflecting number of vehicle-kilometres travelled on different road types); population characteristics (e.g. vehicle ownership, social deprivation index); whether roads were in urban or rural areas; topography; and weather conditions. These were all aggregated at a TLA level. The variables found to be significant when all road types were considered together were:

- Total population;
- Population aged 15-24 years old;
- Percentage of household that drives to work (averaged for TLA);
- New Zealand social deprivation score (explanation follows);
- Number of junctions per kilometre;
- Percentage of traffic travelling on urban roads; and
- Population density.

The social deprivation index was established by New Zealand's Ministry of Health to gauge the affluence/poverty of each census "meshblock" (the unit of analysis used by Statistics New Zealand which generally contains 60 – 100 residents) from the 2001 census data (Statistics New Zealand 2006). The index considers the number of people who:

- are aged 18-59 receiving a means-tested benefit;
- are aged 18-59 and are unemployed;
- have income below a certain threshold when household composition and size is equalised;
- have no access to a telephone;
- have no access to a car;
- are aged less than 60 years and are living in a single parent family;
- are aged 18-59 and have no qualifications;



- live in a household below a certain bedroom to occupancy threshold when household composition and size is equalised; and
- are not living in their own home.

Each category is weighted (listed above in decreasing weighting factors) and a continuous scale is used. An ordinal scale ranging from one (least deprived) to ten (most deprived) is then applied to give a decile ranking system for the whole of New Zealand. (Statistics New Zealand 2006)

Very little indication of a relationship between a TLA's bendiness and its crash rate was found. In most cases, the coefficients of the bendiness measures were negative, indicating that bendiness might be protective. However, only two results were significant at a 5% threshold; the detour ratio and cumulative angle measures used for urban roads.

## **2.3 Implications for this Research**

### 2.3.1 Variables to be Considered

Many previous studies have shown that horizontal curvature is a major contributor to crash occurrence. Generally studies that focus on individual horizontal curves show that safety decreases with increasing radius of curvature.

Obviously there are many other factors that contribute to crash occurrence. Many of these are user related (e.g. speeding, inattentiveness, intoxication), some vehicle related (e.g. faulty brakes, poor steering) and others due to aspects of the road environment apart from bendiness (e.g. vertical curvature, super-elevation at certain locations, sight distance, pavement properties). Thus it was decided that any bendiness or design consistency measures used in this study to explain crash occurrence should be analysed in conjunction with other features that may be present in some areas of the network but not in others.

It would be impossible, however, to obtain data on all the possible influencing variables for the entire road network of New Zealand and too much time and

computing power would be required to examine the effect of each variable and hold them constant to determine the true effect of bendiness. Also, some factors that are known to contribute to crashes, for example driver inattentiveness, are extremely difficult to quantify and a number of factors probably exist but have not been identified by previous research. It can be hoped that some unaccounted factors (for example skid resistance levels and proportions of speeding drivers) may be evenly enough spread across the network so that a large enough sample would effectively hold such factors constant for the analysis. Hence this research was not expected to account for all possible influencing variables.

In terms of choosing which variables to account for, more recent studies suggested that the external features of a horizontal curve, such as its context in the surrounding road network and road environment may be more important than its internal features such as radius of curvature, superelevation etc that were the focus of traditional studies. This implies that this research, in its attempt to analyse road networks (and thus focus on external factors) rather than isolated curves (and their internal factors), is relevant to road safety. Also it was expected that, by considering a large number of crashes each with differing characteristics, the effects of other possible influencing variables would be minimised.

Most studies recognised that, of all possible influencing variables, the variation in traffic flows between two different locations has a large effect on the crash rates and should be accounted for.

### 2.3.2 Methods of Evaluation

One suggested method of considering external features is assessing the “bendiness” (which can be defined in many different ways) of a network or length of road.

Another method of considering external features is to gauge the level of consistency between successive elements. Generally this involves some measure of vehicle speeds (design or operating). It has been shown that good consistency between elements ensures that the task of driving is not too demanding, as drivers are able to form accurate perceptions about the requirements of the road environment. However, it is important that drivers do not find the task too easy so that they remain attentive.

Therefore, it seems appropriate that this research trials variations of both bendiness and consistency methods and also looks for other ways to compare variations between networks.

GIS is a tool that will be helpful in analysing large amounts of data with many spatial aspects or layers. However, there will be some problems associated with the use of data in GIS due to the necessary simplifications and assumptions needed to digitise real world occurrences. The statistical techniques used to determine significant contributing variables should also be carefully chosen. Research suggests that Poisson and negative Binomial models are most appropriate for crash risk models.

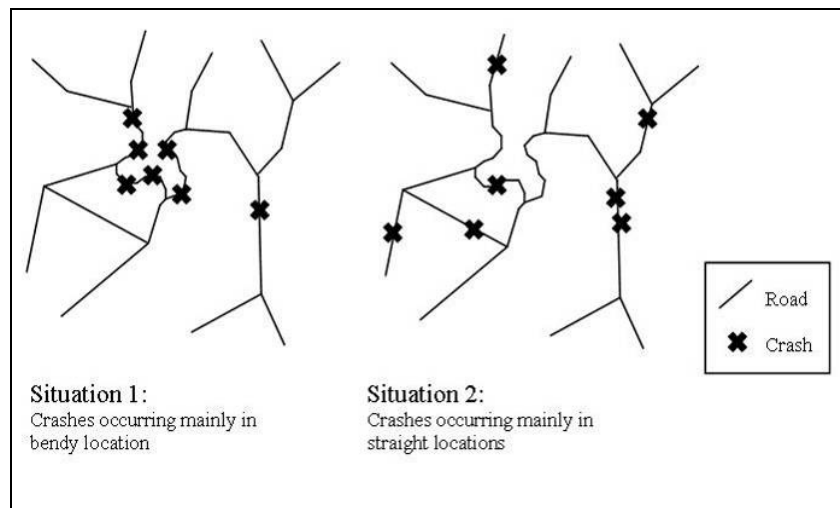
### 2.3.3 Suggested Improvements to Previous Study Methods

One study on the effect of bendiness in New Zealand road networks on crash occurrences has already been performed. The motivating study used a carefully formulated methodology but did not find any strong relationship between bendiness and traffic crashes in New Zealand.

The bendiness measures used in the motivating study gave a comprehensive evaluation of bendiness and were based on many previous studies. Thus it was determined that these measures, with the addition of a measure aimed at representing design consistency, would form the starting point for this research. However, it appeared that several modifications could be made to the study to ensure consistency with the underlying principles of the bendiness measures and that from these changes clearer results may emerge.

The motivating study and Haynes *et al*'s (2007a) British study examined large regions and essentially compared the degree of bendiness to the crash rates for each. It should be noted that the motivating study intended to use an aggregated approach to “break away from a micro-scale focus and allow the possibility of contextual effects” (Haynes *et al.* 2007b). Here “contextual effects” referred to properties of a region such as socio-economic characteristics. However, when using this aggregated approach the motivating study did not attempt to determine whether or not the crashes occurring in a region actually occurred near the bendy locations.

To illustrate this point, suppose that a region had a small number of curves that were all concentrated together in one part of the region. Overall, the region would appear to have a very low level of bendiness and therefore, if it had a large amount of crashes, it would appear that bendiness is protective. However, it may be that all the crashes occurred in the region's only bendy location so the reality would be that bendiness is hazardous but this would not show up in the analysis. This is illustrated in situation 1 of Figure 2.8, whereas situation 2 would indicate that bendiness is no more likely to cause crashes than straight roads.



**Figure 2.8 Concept Showing Potential Pitfalls of Aggregation**

It was anticipated that analysing the roads recently travelled by each driver involved in a crash and computing bendiness measures on a crash by crash basis would give a better idea of the effect of bendiness on crashes than aggregating to a TLA level.

This approach may also improve the estimations for effects of other influencing variables, especially traffic volumes. The motivating study accounted for volumes by taking average traffic flows over TLA regions. This may be too generalised, as within a TLA there can be large variations in flows, especially those with clearly defined road hierarchies.

Ideally the route of each driver would be known and therefore analysis of the roads actually travelled leading up to each crash that influenced the driver's perceptions would be possible. However, the New Zealand Crash Analysis System (CAS), which

provided the crash data, does not provide much information on the routes taken by drivers. A general indication of the direction of travel (north, south, east or west) is given but this can be ambiguous and inconsistent, for example a crash occurring on winding route to a northern destination may occur at a curve where the direction of travel is actually south bound. In this situation the direction of travel could be coded as either north or south, depending on the reporting officer.

It is also possible to access the crash reports filled in by attending officers which are generally accompanied by diagrams that illustrate the vehicle's trajectory. This could clear up any confusion from the coded directions but would be a very time-intensive process. Also, even if the direction of travel were known there is no way of distinguishing between possible routes heading in the same direction that lead up to the crash location.

CAS does record the home addresses of drivers (although the general public cannot normally access this) and it could have been possible therefore to infer a route of travel by assuming that all trips originated from home. This was considered too big an assumption to make, as people partake in many activities (and hence have many origins and destinations) over the course of a day. In addition to this, the extraction of drivers' addresses from the database and subsequent geo-location would be a highly time-intensive task.

Multi-vehicle crashes that involve more than one direction of travel (for example head-on crashes) would complicate this process even further as a distinction would have to be made regarding whether the roads travelled by one particular driver or both drivers had influenced the outcome of the crash.

Thus it was assumed that a proxy measure of using "influence areas" of a certain distance around the crash location and analysing all possible routes from the crash location to the edge of the zone would be sufficient. It was assumed that this area could be based on either travel distance or time. It was intended that alternative routes would be assigned a probability that they were the route taken, according to traffic flows or other variables.

Previous research on New Zealand roads has shown that the urban and rural situations are very different due to factors such as network function, land use environment and presence of pedestrians and cyclists (Appleton *et al.* 2006). The motivating study and Haynes *et al.*'s (2007a) British study attempted to distinguish between urban and rural areas by using variables such as the percent of road in the region that is urban and holding these factors constant in the analysis. They also analysed different road types separately. However, this research assumed that it might be more appropriate to develop different models for the two cases as they involve very different traffic characteristics and that bendiness may not actually be a major contributor to urban crash occurrence.

Urban areas are characterised by more closely spaced intersections, slower speeds, higher traffic volumes and different distractions than rural areas. Thus, it was also assumed that the bendiness measures used in the motivating studies may not be transferable to a crash-by-crash analysis for urban areas due to the number of possible routes between two nodes. Likewise it was assumed that crashes that occurred at intersections and those that occurred at mid-block locations would also have different characteristics and might require different models and bendiness measures to be developed.

It was assumed that the initial use of one model for all cases would highlight the types of distinctions that should be made before the data could be subdivided and separate models developed.

### **3.0 General Method**

This section outlines the general approach to measuring the effect of bendiness on crash occurrence formulated by this research. The main assumption of the method, as outlined in the previous section, is that crashes should be analysed individually, rather than at an aggregated level, with weightings applied to all possible routes surrounding a crash.

The first basic step of the method is to process crash and road network data into formats compatible with the calculation methods. Next calculations of bendiness (according to several different definitions) should be performed to give bendiness values for each crash location. These results can then be used to model and predict the effects of road network bendiness on traffic crash occurrence.

#### **3.1 Data Preparation**

The first step required is to obtain the necessary data and process these into the formats required for calculation of bendiness measures. A suggested method of data preparation is shown in Figure 3.1.

It should be noted that this process was tailored to suit the specific requirements of the bendiness measures used in the case study. Thus a concept of the desired outputs, subsequent steps and a general idea of how those outputs could be obtained was required in the first step of data preparation. While similar studies may have different desired outputs and work with different data sets, some general requirements of the data can be suggested:

- Data must be available for the crashes and road network of the study region. It is desirable to have flow data corresponding to the road network so that measures can be weighted according to vehicle exposure, but alternative methods can be developed if no flow data are available;
- The crash data must have location information. Generally this will be available as geographic coordinates which can then be referenced to the closest link of the road network. Some crash data may come from a source

that uses a route positioning system; this would be useful only if a corresponding road network dataset is available as it would be time intensive to add route position information to a road network;

- It would be desirable (but not imperative) to have other information about crashes, for example the severity, date of occurrence, collision type, contributing factors and road conditions;
- The road data must correspond to a known map projection (preferably the same as that of the crash data). It must contain centreline position but could also have more detail (e.g. edge-lines). It must be of a format that can be converted into a collection of straight lines and vertices;
- If separate road and flow data sets are used it is important that there is a means of linking the two. Thus the flow data should include either spatial references or an ID field that corresponds to one of the road data fields; and
- If comparison between different regions is intended, an additional data set of regional boundaries is required. Alternatively, if the road data already has regional information this can also be assigned to the crash data and no further data is required.



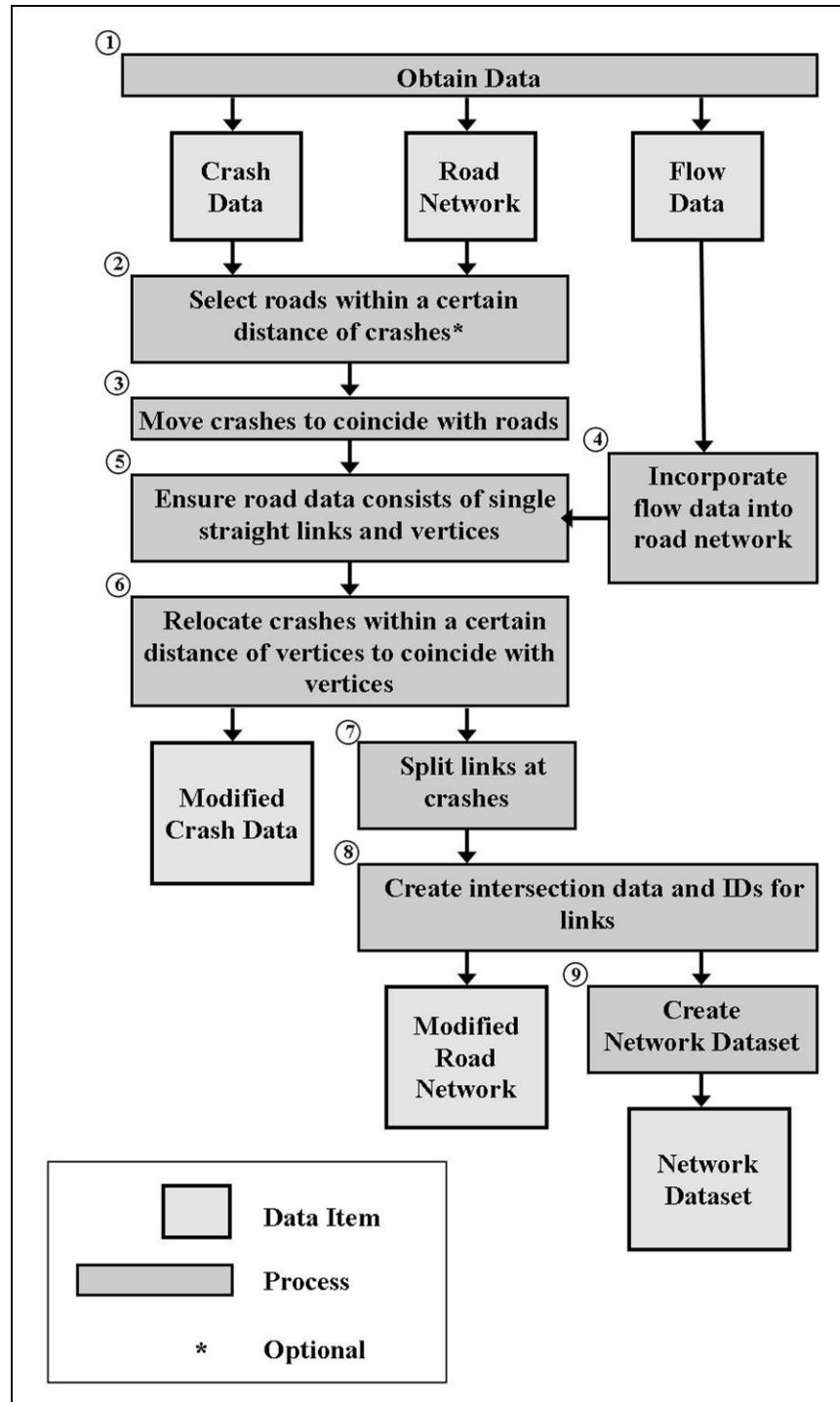


Figure 3.1 Data Preparation Flow Diagram

A general explanation of the steps outlined in Figure 3.1 follows; specific details of the method applied in the case study are found in subsequent sections of this report.

### **1. Obtain data**

As different types of data are required, it is likely that they will each have to be obtained from different sources. This step does not require much technical expertise, except for knowledge of the requirements of the data and an idea of what operations will be performed on them.

### **2. Select roads within a certain distance of crashes (optional)**

If the analysis is to be done only on the roads surrounding crashes, the size of the sample to be processed (and hence computing time) can be significantly reduced by initially selecting only the roads within the buffer distance of interest. The size of this buffer,  $\kappa$ , is a parameter to be set by the analyst. It may be preferable to use a large buffer and later reduce the sample further if it is decided a smaller area would be appropriate. This step assumes that the crashes are sparsely distributed; if not it may be worth omitting this step as it will not result in much of the network being removed. It should also be considered whether or not analysis of roads where no crashes have occurred should be performed for comparison. If so, this step should not be included and the whole road network should be used.

### **3. Move crashes to coincide with roads**

The bendiness measures outlined in the following section require that the crashes intersect the road network lines. Due to the data's nature (e.g. using widthless lines to represent roads) and inaccuracies (as a result of digitising etc) it is likely that most crashes will not perfectly coincide with the roads on which they occurred. Thus some operation must be performed to relocate the crashes so that they coincide with the nearest road links.

### **4. Incorporate flow data into road network**

A field for traffic flow on each link should be created in the road data and somehow (either through spatial or attribute data) the flow data should be joined to the road data.

### **5. Ensure road data consists of single straight links and vertices**

If the original road data consisted of polylines, these must be divided into individual straight lines, each with a unique identification number and coordinate information for both of its ends. Vertices should be located at all ends of the straight lines.

### **6. Relocated crashes within a certain distance of vertices to coincide with vertices**

Problems, such as incorrect link selections, can result when very small links are created from the splitting of links at crash points. To avoid such problems, crashes that are close to vertices should be relocated to coincide with these vertices. This requires another parameter to be determined by the analyst. A reasonable suggested value for  $\beta$ , the minimum allowable distance between a crash and vertex, is 1 metre.

### **7. Split links at crashes**

All road links that have crashes situated partway along their length should be divided at the crash location so two new links are created. New links should also have references for their start and end vertices.

### **8. Create intersection data and IDs for links**

Each link should have information regarding how many other links it joins at either end. This can be done by finding for each vertex how many road links it intersects and joining this information back to the road data through the references created previously. An identification field should also be established so that each individual road link has a unique number.

### **9. Create network dataset**

The bendiness measures require the use of a “network dataset” (used to model the connectivity of roads) so that areas of influence and route paths can be determined. Depending on the type of GIS software used this may not be required or may have a different name.

Figure 3.2 shows a concept of a section of the processed road network and crash data where all links that lie within an  $\alpha$ km buffer zone of the relocated crash are included.

(The vertex created at the relocated crash location has also been classed as a node for future calculations.)

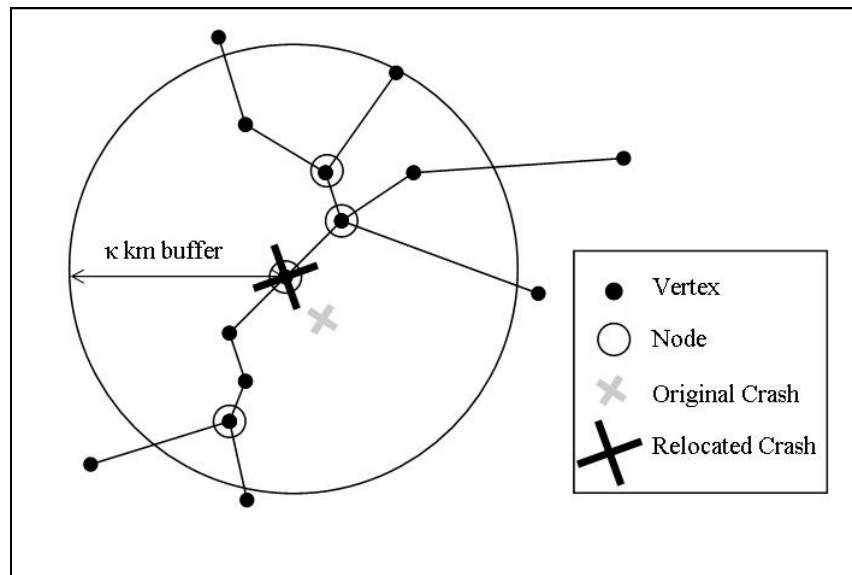


Figure 3.2 Processed Road Network Concept Diagram

### 3.2 Calculations of Bendiness Measures

The second step involves the calculation of bendiness measures for each of the crashes in the modified crash data. This process, detailed in Figure 3.3, uses the links and vertices from the modified road network and also incorporates the network dataset for use in the area of influence and route functions.

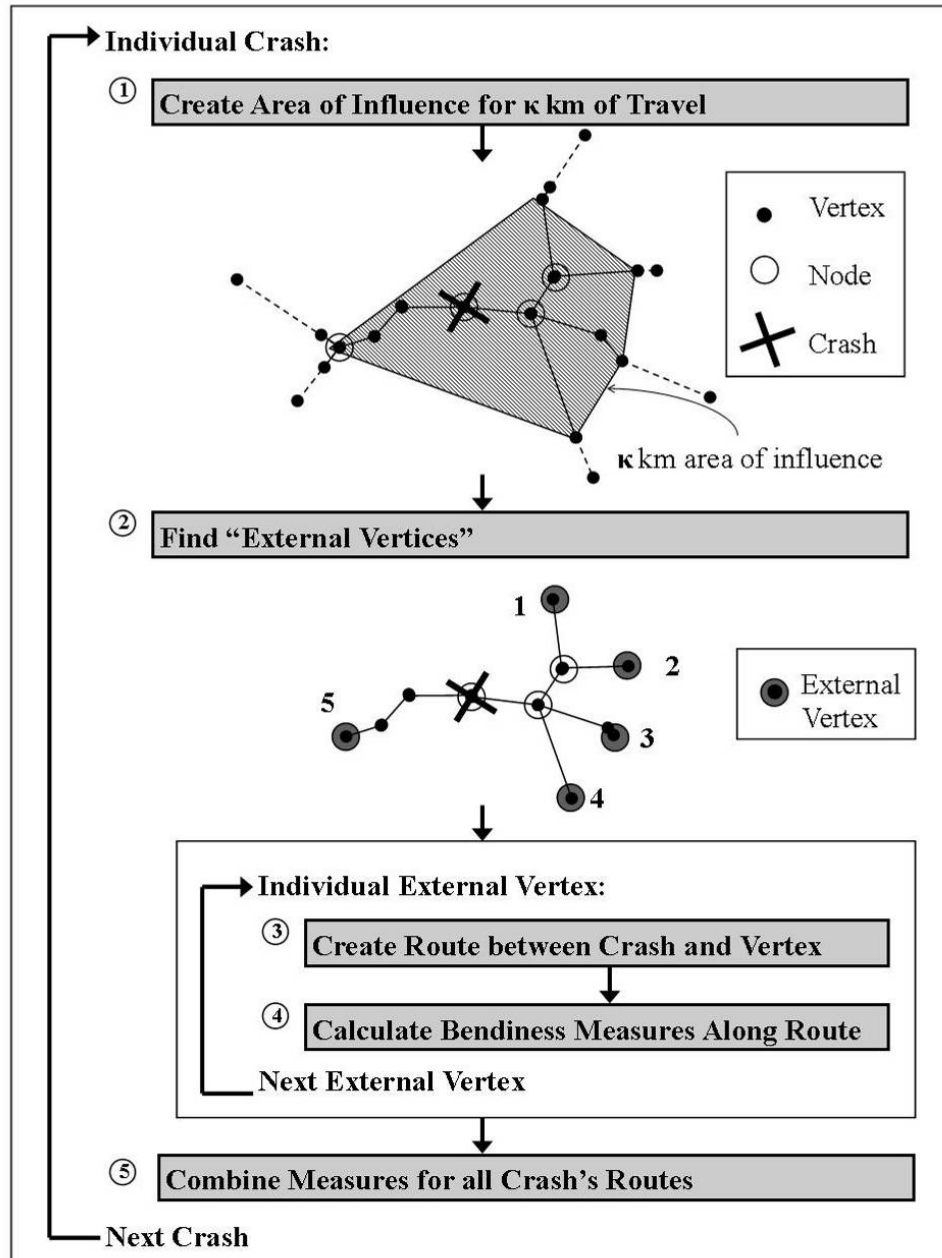


Figure 3.3 Bendiness Measures Calculations Flow Diagram

### 1. Create area of influence for $\kappa$ km of travel

The area of influence is the area that encompasses all points within in a travel distance of  $\kappa$ km. Here the value for  $\alpha$  used is the same as for the data processing step but the area of influence is not equal to the buffer area as the latter encompasses all points of the road within a radial distance of  $\kappa$ km. The reason for using the radial distance

initially was that it is a much faster operation to perform and can speed up the process of finding in the influence area by decreasing the size of the network considered.

## 2. Find external vertices

The external vertices are those immediately outside of the area of influence. They can be found by selecting the links that partially touch the area of influence and then selecting the vertices that touch these links but do not touch the area of influence.

## 3. Create route between crash and vertex

For each external vertex a route must be created between the vertex and the corresponding crash. This route should be the shortest route possible between the two points as this is the most logical solution that does not involve having to model the traffic flows over the whole network. Links corresponding to this route should be selected and ordered according to their position along the route so that calculations can be performed.

## 4. Calculate bendiness measures along route

Using the geometric properties of the ordered route links the bendiness measures can be calculated. This requires the use of programmed calculations, the format of which will depend on the software employed.

## 5. Combine measures for all crashes' routes

When all routes have been considered the overall bendiness measures for the crash should be computed. A variety of different methods of combining the values of each route could be considered:

### 3.2.1 Method 1: A simple summation of each route's bendiness values

This method is the most simple one possible as the bendiness measures of all routes are added together without any weighting factors applied, as shown in Equation 16:

$$BM_{Crashj} = \sum_{r=1}^R BM_{Routejr}$$

**Equation 16 Weighting Method 1**

Where:  $BM$  = type of bendiness measure (detour ratio, mean angle etc)

$Crash_j$  = one particular crash

$Route_{jr}$  = a route belonging to the influence area of crash<sub>j</sub>

$R$  = the total number of  $route_{jr}$ s in the influence area.

This method is not recommended as it does not involve any weighting of different possible routes and therefore assumes that all routes are equally likely to be travelled along. Areas with many minor side roads would be over-estimated by this method as a greater number of possible routes (regardless of the low probability that the side roads would be used often) would produce higher bendiness values.

### 3.2.2 Method 2: Weighted according to the number of routes in the area

This is considered a slight improvement on the previous method as use of Equation 17 removes the bias towards areas with more possible routes by dividing the value obtained by the number of routes in the area.

$$BM_{Crashj} = \frac{1}{R} \sum_{r=1}^R BM_{Routejr}$$

**Equation 17 Weighting Method 2**

However this method is still undesirable as it does not account for the fact that some routes are more likely to be used than others.

### 3.2.3 Method 3: Weighted according to number of routes at each intersection

This method, mathematically represented by Equation 18, weights routes according to the number of routes available and should be used in situations where no link flow or road hierarchy data are available.

$$BM_{Crashj} = \sum_{i=1}^R \left[ BM_{Routejr} \times \frac{1}{X_1} \prod_{n=2}^{N-1} \frac{1}{X_n - 1} \right]$$

**Equation 18 Weighting Method 3**

Where:  $X_n$  = the number of links joining to a node

$N$  = the total number of nodes along a route (first node = crash position, last node = influence area edge)

#### 3.2.4 Method 4: Weighted according to routes' link flows

This method is recommended given that Equation 19 makes use of flow characteristics along each route and hence best approximates the relative likelihood of each route being used.

$$BM_{Crashj} = \sum_{r=1}^R \left[ \frac{BM_{Routejr}}{\sum_{l=1}^L \frac{d_l}{f_l}} \right] / \sum_{r=1}^R \left[ \sum_{l=1}^L \frac{d_l}{f_l} \right]^{-1}$$

**Equation 19 Weighting Method 4**

Where:  $f_l$  = the flow along link  $l$

$d_l$  = the length of link  $l$

$L$  = the total number of links along a route

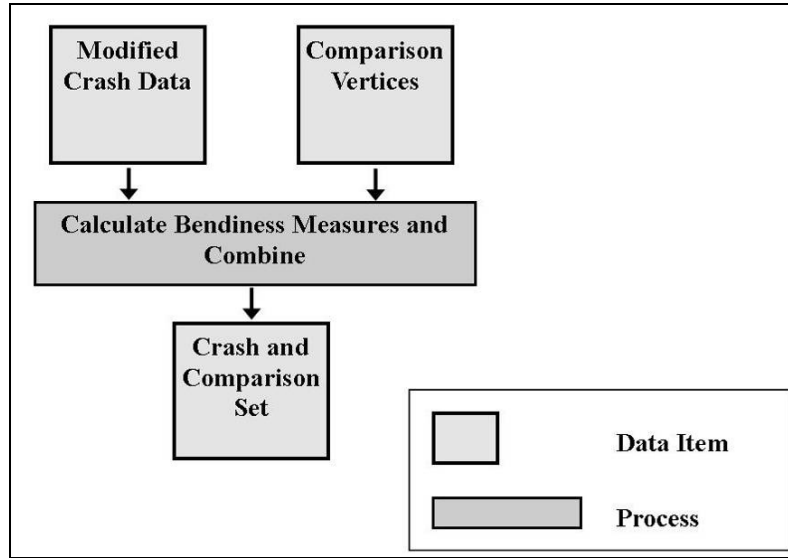
### **3.3 Analysis Methods**

Once bendiness measures have been computed for each crash in the crash sample the values should be combined in some way that indicates the effect of bendiness on crash occurrence. This process depends on the desired output format; it could involve simply taking averages of the bendiness values for crash and comparison sites or by weighting these values according to traffic flows in the vicinity of crashes. Statistical analyses could also be performed to account for the effects of other possible influencing variables that are not related to bendiness such as road characteristics, weather or topography where data is available.

#### 3.3.1 Comparison Data

Rather than simply analysing the crash locations, a more powerful statistical analysis would involve the use of comparison sites. Comparison data should consist of a sample of vertices from the parts of the road network where no crashes have occurred. Bendiness measures and other influencing variables should be obtained for these comparison locations in the same way that they were obtained for the crash locations. Thus a new data set consisting of crash and non-crash locations should be created, as illustrated in Figure 3.4.





**Figure 3.4 Data for Statistical Analysis**

### 3.3.2 Choice of Statistical Model

Each of the data points for this method is classed according to one of two possible outcomes; crash location or not a crash location. This differs to the motivating study which gave, for each TLA, a total crash count and thus had a wide range of possible outcomes. For this method, the mean and standard deviation over all the sites (required to choose between negative binomial or Poisson regression) would simply depend on the ratio of crash to non-crash sites chosen. The model type used by the motivating study is therefore not suited to this method.

Thus, for this method it is suggested that a binary logistic (or “logit”) regression model be used for the statistical analysis. The binary logistic regression model is based on data of two possible outcomes and several influencing variables, and has the form:

$$\log it[Y] = \log\left(\frac{Y}{1-Y}\right) = \beta_0 + \beta_1 B_1 + \dots + \beta_f B_f + \dots + \beta_F B_F$$

**Equation 20 General form of Binomial Logistic Regression (from (Agresti 1996))**

where:  $Y$  = the probability of a crash occurring ( $0 \leq Y \leq 1$ )

$B_f$  = influencing factor (for  $f = 1$  to  $F$ )

$\beta_f$  = coefficient applied to factor  $B_f$

When applied in this method, the influencing factors should include different bendiness measures as well as other non-bendiness related factors.

Several indicators are worth noting when assessing the performance of a binary logistic regression model. The “p-values” calculated for each coefficient test the hypothesis that the coefficients are not equal to zero; hence a p-value lower than the specified  $\alpha$  level (degree of certainty) is desired to include that coefficient. In addition to this, a test that all coefficients are equal to zero can be performed; this gives a “G-value” and its associated p-values, both of which should be as low as possible. (Minitab Inc 2003)

Many goodness-of-fit tests that can be applied to binary logistic regression models exist, the most common of these are the Pearson and deviance tests. These tests both give Chi-squared values, which should be as low as possible, and p-values, which should be above the specified  $\alpha$  level so that the null hypothesis of the model being a good fit cannot be reasonably rejected. The Pearson test is more suited when expected frequencies average between one to ten. (Agresti and Finlay 1997) Therefore, since the samples analysed in this research will involve hundreds of data points for each crash and non-crash outcome the deviance measure will be used to assess the goodness of fit.

Concordance measures can also be calculated. This involves comparing a model’s predictions for each combination of pairs of different outcomes. A concordant pair occurs when the probabilities predicted are consistent with the observed outcomes (e.g. of the pair, if the crash location was assigned a higher probability of crash occurrence by the model than the non-crash location the pair would be termed concordant). A discordant pair occurs when the probabilities predicted are opposite to that of the observed outcomes. A tied pair occurs when the probabilities of both outcome occurring are predicted to be equal. Hence, the higher the percentage of concordant pairs, the more accurate the model. (Minitab Inc 2003)

## 4.0 Case Study

This section introduces the case study region and data to which the general method presented in the previous section was applied.

### 4.1 General Description of New Zealand

New Zealand, shown in Figure 4.1, is a country in the south-west Pacific region. New Zealand is comprised of two main islands (North and South) and several smaller islands. Its 270,500 square-kilometres, a similar area to that of Great Britain or Japan, are home to almost 4.2 million people of varied cultures, making New Zealand one of the least crowded countries in the world (Statistics New Zealand 2007).

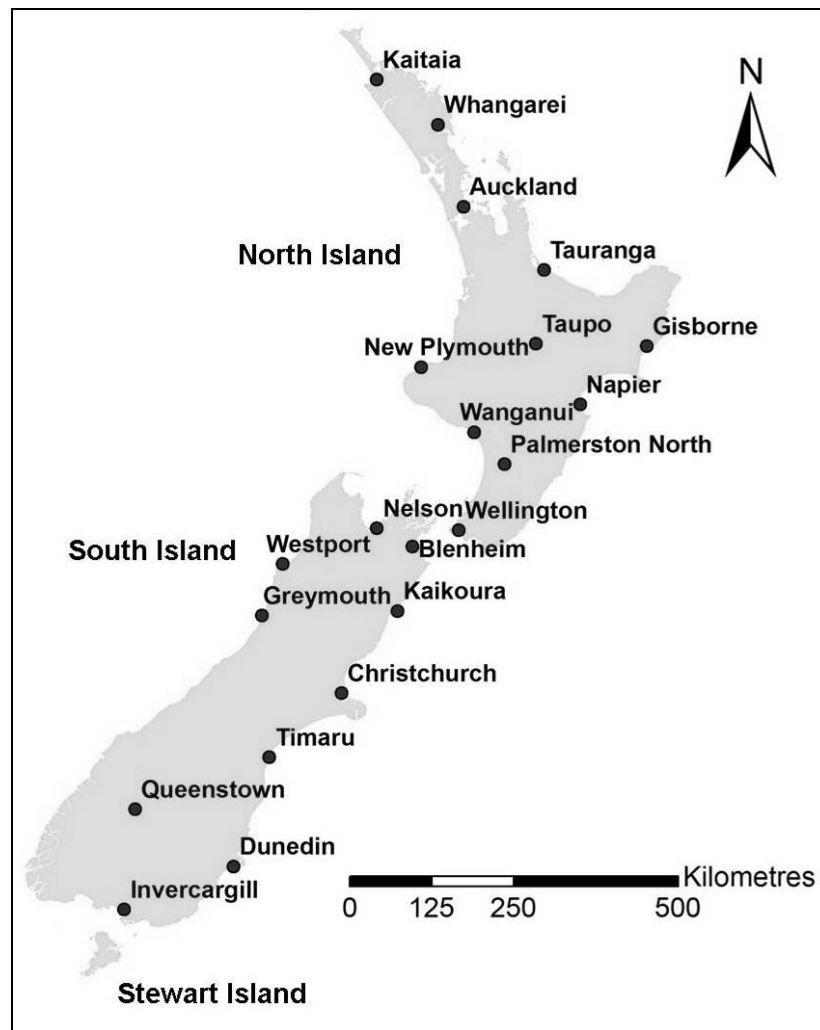


Figure 4.1 Map of New Zealand (adapted from LINZ (2005))

New Zealand is divided into 73 Territorial Local Authorities (TLAs), each of these having a Road Controlling Authority (RCA) responsible for control and maintenance of local roads within the TLA. The country's state highway network, controlled by Transit New Zealand, consists of almost 11,000km of road and carries approximately 50 percent of the country's total traffic (Transit New Zealand 2007). As of 2004, there was an average of 1.8 cars per New Zealand household (Ministry of Transport 2004) with 9,992 injury and 375 fatal crashes occurring (Ministry of Transport 2006).

#### 4.2 Data Relating to Bendiness

In order to assess the effect of road bendiness on traffic crashes in New Zealand, two fundamental data sets were required as detailed below in Table 4.1. Unless otherwise stated, all data used was projected according to the 1949 New Zealand Map Grid coordinate system.

**Table 4.1 Principal Data Sets and Sources**

<b>Data</b>	<b>Description</b>	<b>Source</b>
Crash Records	Fatalities on NZ roads from 01/01/96 to 31/12/05	CAS (Ministry of Transport 2007)
Road Network Data	Centreline data for New Zealand's road network	University of Canterbury Geography Department via Terralink International

Each data set required different modifications before bendiness measures could be calculated. While it was desired that all necessary modifications would be identified from small trials before the bulk of the calculations were done this was not the case. During the calculation of bendiness measures (as described in Section 5) some anomalies in the data sets emerged that required additional processing of the data. Also, some anticipated data sources (such as flow data for all of the country's roads) were not obtained, which changed the requirements of the original data. Thus some modifications to the data explained in this section were actually made in reaction to challenges encountered in the course of investigation. It is important to note that, had these challenges been known before calculations began, it is likely that some of the modifications would have been made differently.

#### 4.2.1 Description of Crash Data

The crash data were obtained via Land Transport New Zealand's Crash Analysis System (CAS) by searching nationwide for fatal crashes (i.e. traffic crashes resulting in the death of one or more people) that occurred in the period from 1 January 1996 to 31 December 2005. 4056 fatal crashes were identified.

The selection criteria for the crash data were made primarily so that it would be the same as that used in the motivating study, but it is worth mentioning why Haynes *et al*'s (2007b) used such criteria. Firstly, only fatal crashes were used as New Zealand is considered to have a poor reporting rate for crashes of moderate to low severity, with high variances in reporting rates between TLAs (Alsop and Langley 2001).

For such studies, periods of five to ten years are generally recommended (LTNZ 2004). As fatal crashes are much less frequent than crashes of lower severity a ten-year time period was selected to reach a sample set large enough to draw statistically significant conclusions. It should be acknowledged that over a ten-year period significant changes may have been made to the road network (and other influencing variables could have changed also). It is also worth acknowledging that approximately 10,000 injury crashes are reported in New Zealand yearly (Ministry of Transport 2006). So, even though use of injury crashes may not have been as reliable as fatal crashes in terms of reporting rates, it would have allowed a shorter analysis period and thus increased reliability in terms of consistency between the roads the crashes actually occurred on and the road network used in the analysis. These points outline some of the trade-offs that must be accepted when working with such data.

Along with a unique identification number and spatial coordinates (according to the 1949 New Zealand Map Grid reference system), each crash listing contained information regarding:

- Temporal characteristics – date, day and time of crash etc;
- Intersection characteristics – whether the crash occurred at an intersection and, if so, the nature of its control;
- Crash movement type and causative factors;

- Geometric characteristics – including an indication of whether the road was straight or had an easy, moderate or severe curve;
- Environmental characteristics – weather, lighting etc;
- Road environment characteristics – speed limit;
- User characteristics – age of pedestrians or cyclists; and
- Severity of crash – number of people suffering fatality or serious or minor injuries.

The 4056 fatal crashes caused a total of 4684 fatalities (some crashes involved multiple fatalities) and also caused 1924 severe injuries and 2289 minor injuries (note that other severe and minor crashes have also occurred in the region over the study period, the figures here are only those injuries associated with fatal crashes). The percentages of these injuries compared to whether they occurred on straight or curved sections of road are very similar to those of the fatalities.

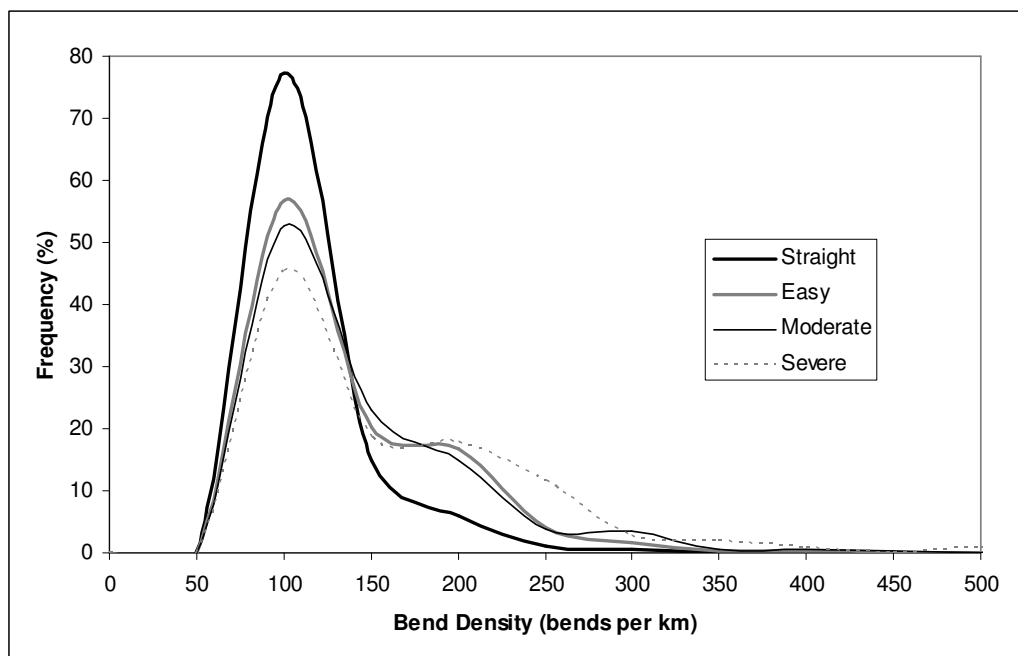
While the CAS data on geometric characteristics of crash sites gives some indication of bendiness CAS does not contain any information on the number of straight and curved sections and the traffic flows over the whole of New Zealand's road network (and hence exposure to road elements). Thus the crash data does not allow any inference of the effect of curvature on crash rates. However, initial breakdown of the data did reveal that about 50% of crashes were recorded to occur on a curve of some description, as shown in Table 4.2. This is similar to Lamm *et al*'s (1992) estimate for American roads.

**Table 4.2 New Zealand Crashes on Curves (1996-2005)**

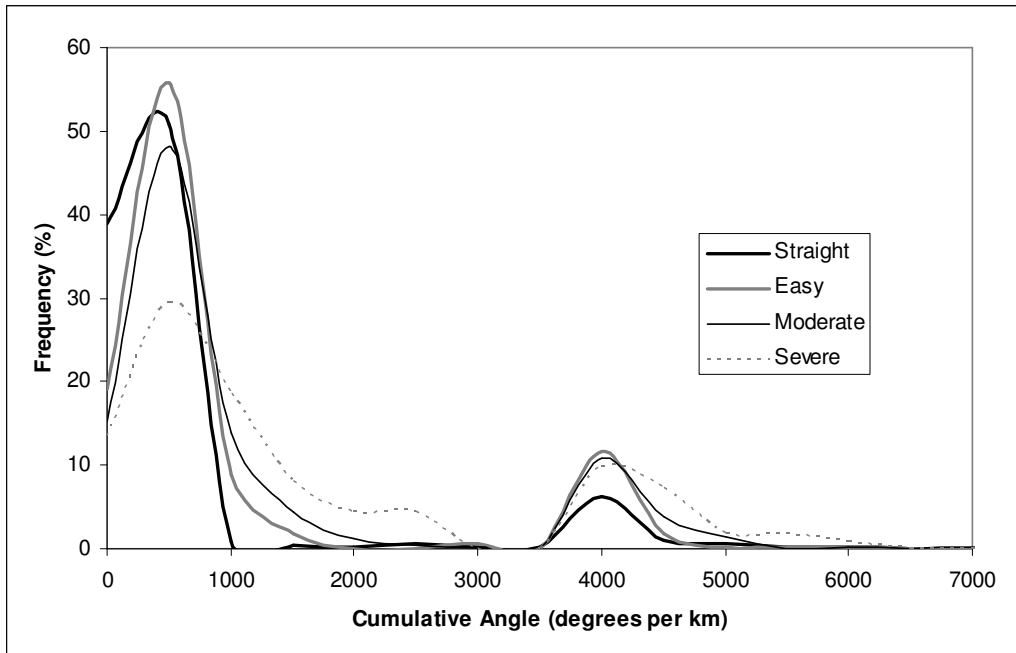
	Number of fatal crashes	Percent of fatal crashes	Number of fatalities	Percent of fatalities
Straight Road	2032	50.1	2289	48.9
Easy Curve	889	21.9	1073	22.9
Moderate Curve	992	24.5	1167	24.9
Severe Curve	143	3.5	155	3.3
Total	4056	100.0	4684	100.0

While CAS records some information regarding the curvature of roads, it should be noted that the classification relies on a subjective appraisal made by reporting officers and does not have any strict quantitative guide in determining the degree of curvature. A simple comparison of the CAS curve types and GIS-computed bendiness measures was performed. Figures 4.2 and 4.3 show the frequency distributions for the bend density and cumulative angle measures calculated for influence regions 25m either side of each crash.

Only crashes on stretches of road that had no intersections over the 50m span were considered, to avoid the confusion of multiple possible routes. Thus it was not appropriate to evaluate the detour ratio over the 50m spans, as the detour ratio was concerned with straight distances between intersections. This was unfortunate, considering that the detour ratio had been identified as the most appropriate measure for the main research, but other measures still gave an indication of the consistency of the CAS curve type classification.



**Figure 4.2 Bend Density Frequency for different CAS Curve Types**



**Figure 4.3 Cumulative Angle Frequency for different CAS Curve Types**

From Figure 4.2 it can be seen that there are no specific ranges of bend density that correspond to the different CAS curve types. All of the distributions peak at the same point (about 100 bends per km); this is probably an indication of the resolution of the data (as “bends” here are defined as vertices). However, in the higher regions of bend density, there is a definite trend. Stretches with  $BD > 150$  were unlikely to be classified as straight roads and severe curves were more frequent at the higher ends of the scale. There was little distinction between easy and moderate curves; this suggests that the CAS classifications are too subjective for these purposes.

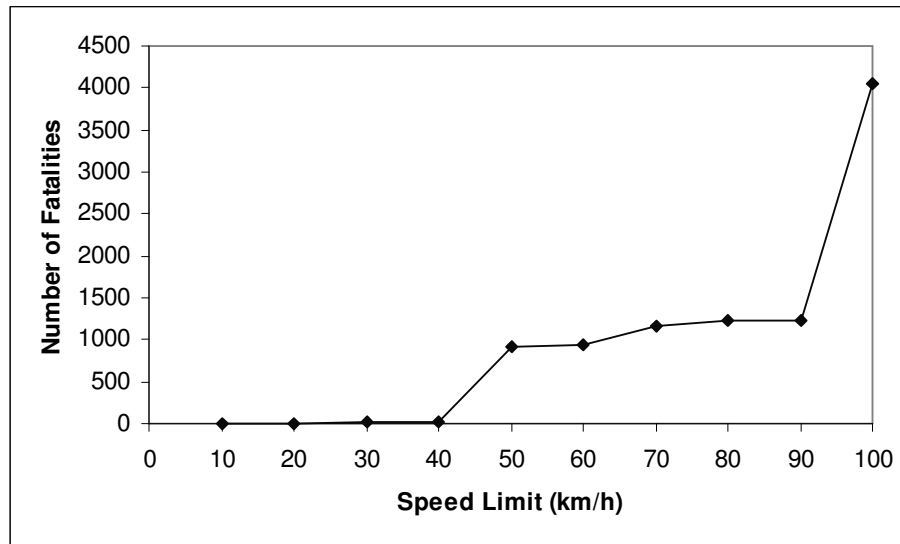
Similar trends are displayed in Figure 4.3, the cumulative angle frequency distributions. It seems strange that there are few locations in the  $1500 < CA < 3500$  degrees per kilometre range. Closer inspection of the crash locations reveals no obvious reason explaining the peak that occurs at approximately  $CA = 4000$  degrees per kilometre. Crashes in this range were spatially distributed across the country similarly to all other crashes. They generally appeared on curved sections with closely spaced vertices, whereas the remaining crashes are situated near one or more long straight links. Thus the peak at 4000 degrees per kilometre seems to be a consequence of resolution; shorter links are required to digitally represent curved sections of road than are required for straight sections.



Given this, it seems odd that a significant number of road sections with high cumulative angle values were recorded as straight or easy-curve in CAS crash reports and, conversely, a significant number of sections with very low cumulative angle values were recorded as having moderate or severe curves.. This is probably partly due to the subjectivity of the CAS classifications and partly due to incompatibilities between the CAS and bendiness measure methods. The bendiness measures were calculated for a set distance whereas the CAS classifications are more aimed at whole curves, the length of which may be significantly greater or smaller than the 50m influence area used. Also, the reporting officer knew the direction of travel and based their assessment on the portion of road that had been driven on, whereas the bendiness measures were based on both directions from the crash.

Thus, this preliminary investigation cannot be used to make any strong conclusions regarding either the appropriateness of the CAS classifications or the bendiness measures, but it is useful to see that the rankings of the two are generally consistent. For example, at lower bendiness values straight roads are most common, followed by easy, moderate and severe curves, and for higher values of bendiness this order reverses.

Other general characteristics of the fatal crashes studied were observed. A summary of the crash movement types involved is given in Appendix 1, the predominant movement types were head on, loss of control and cornering. Cars were by far the most common vehicle type involved, followed by vans/SUVs, trucks and motorcycles. As shown in Figure 4.4, most of the crashes occurred in areas where the speed limit was 100km/h; this was expected given that severity increases with speed (ACC and LTSA 2000). It is important to note that the speed limit of an area is not always indicative of the recommended safe travelling speed, especially in regions of high bendiness.



**Figure 4.4 Cumulative Frequency of Speed Limit versus Fatalities**

Information was obtained on whether or not each crash occurred “at” an intersection. “At” an intersection is defined as being within 10m of the intersection’s centre (Ministry of Transport 2007). Based on this, the number of mid-block (i.e. non-intersection) crashes was significantly greater than the number of intersection crashes. However, information regarding the junction type and control type (signals, roundabout etc) of the intersection was also supplied for each crash. The number of crashes identified as being at an intersection was lower than the number of crashes with a junction type specified and significantly lower than the number of crashes with a control form specified. (For example crash number 9620063 in central Christchurch had cross roads specified as the junction type but was not recorded as occurring at an intersection.) This indicates either a poor understanding of the reporting requirements (e.g. the definition of an intersection crash) or significant data errors. For the purposes of this research, it was assumed that the intersection/mid-block field would be the most correct, as it is the least open to subjective interpretation.

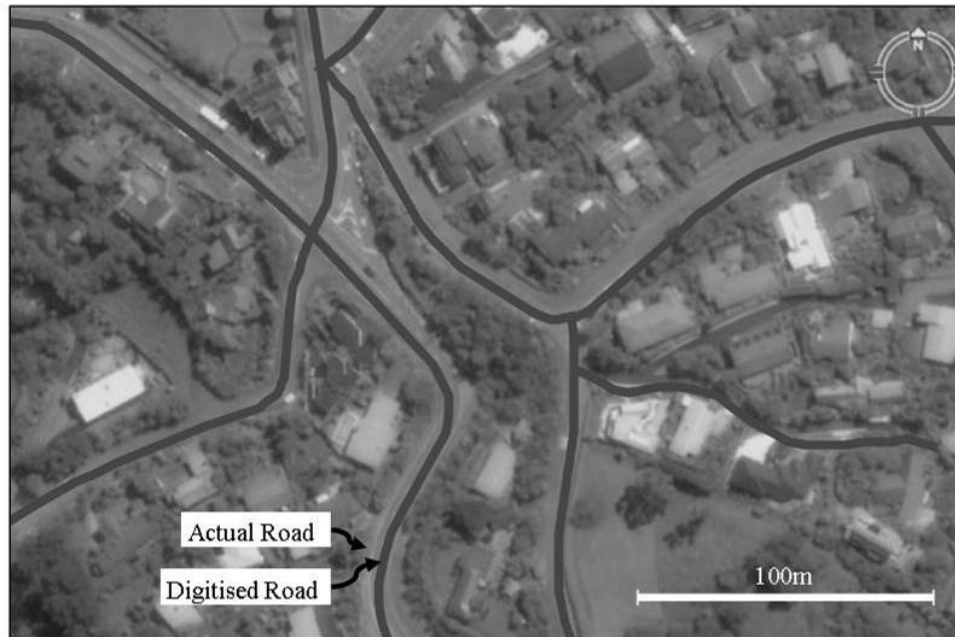
#### 4.2.2 Description of Road Network Data

The road network data used in the GIS analysis was the same data as used in the motivating study. The network data consisted of centreline location information compiled from several different sources and was projected using the 1949 New Zealand Map Grid coordinate system. Each arc (i.e. a series of straight links between

nodes) was stored as a polyline with coordinates for the start and end nodes and a unique identification number. Also, each arc was accompanied by information on:

- Length;
- Dates of creation and modification;
- Road name;
- Road type;
- Suburbs at either end of the road;
- Place (TLA) name;
- Road class (minor rural, major rural, minor urban, urban arterial, motorway, state highway);
- Rural/Urban index (according to the road classification);
- Surface type;
- Number of lanes; and
- State Highway number (if applicable).

The road data did not represent physical roads exactly due to the scale of digitisation and the fact that the lines were based on centreline positions only. For example Figure 4.5 shows a satellite image of a section of Christchurch's road network, in the suburb of Cashmere. The corresponding section of the digitised road network has been overlaid on the satellite image, with the lines (which are generally assumed to be without width in the GIS) thickened to aid the visualisation.



**Figure 4.5 Digitised Road Network Compared to Satellite Image from (Google Earth 2007)**

It can be seen in Figure 4.5 that the digitised road network does not always correspond exactly with the centre of the road; this is partly due to the use of straight lines and partly due to the digitisation process. Also, the lines do not represent the complexity of the intersections which have different lane configurations and widths than the mid-block sections.

It can be assumed however that the level of accuracy is reasonably consistent for all roads in the network. Hence, as this research is concerned with network effects, not the effects of individual elements, the road network data should be sufficient to give reliable results.

#### 4.2.3 Modifications made to Crash Data

It was important to consider the relationship between the crashes and road network. While the crashes all had coordinates obtained by global positioning systems (GPS) they did not always coincide exactly with the road network, as can be seen in Figure 4.5 which details the crash-to-road distance for the entire crash sample. Such occurrences were in part due to the inaccuracies of both the crash coordinates (poor use of GPS, recording errors etc) and the road network data (i.e. the fact that it

consisted of straight lines to represent the actual road network and was created by the process of digitising). Also, the fact that the road network data used lines (objects with length but assumed to have no physical width) to represent the centreline positions of actual roads, meant that unless a crash occurred exactly on the centreline, regardless of the accuracy of crash and road coordinates, it would not be exactly coincident with the digitised road.

As can be seen in Figure 4.6, the vast majority of crashes were within 10m of the road centreline data, an acceptable difference considering standard road widths and expected inaccuracies. However, some crashes were recorded at great distances from the road. Closer inspection showed that these crashes were generally associated with off-road activity and occurred in regions such as Canterbury's Waimakariri riverbed where four-wheel driving activities are common. Such crashes were discarded from the set. This reduced the sample size from 4056 to 4019 crashes.



**Figure 4.6 Cumulative Distribution of Distance between Crashes and Road**

To ensure compatibility between the crash and road data, the crashes were relocated to the nearest point on the road network. Crashes within one metre of a vertex were

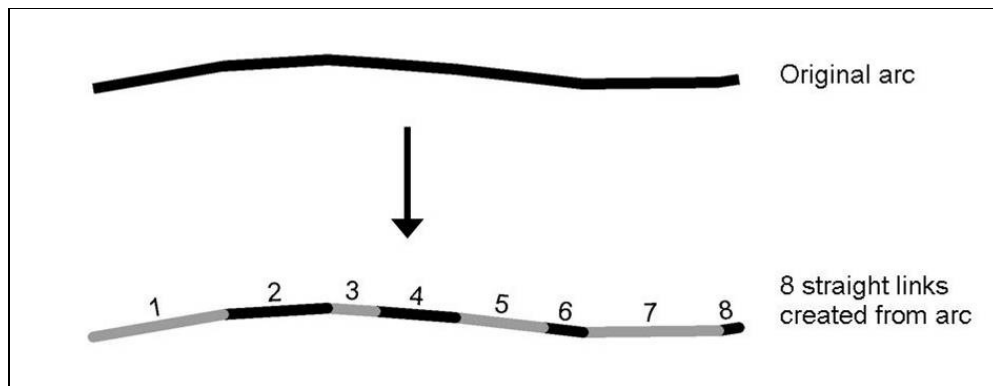
relocated to the vertex to avoid complications that would result from crashes being very close to but not coincident with intersections.

Crashes were also subsequently classed as being either rural state highway (1509 crashes), rural non-state highway (685 crashes) or urban (1825 crashes).

#### 4.2.4 Modifications made to Road Network Data

The road network data were initially represented spatially as arcs rather than individual links. This meant that information was not provided for the start and end coordinates of each individual link and made initial analysis of the bendiness measures difficult. Different methods and computer packages were trialled to find the most appropriate technique. The chosen technique involved using the specialised ArcInfo extensions for ArcMap but is presented here in a generalised form that should be applicable to any GIS software.

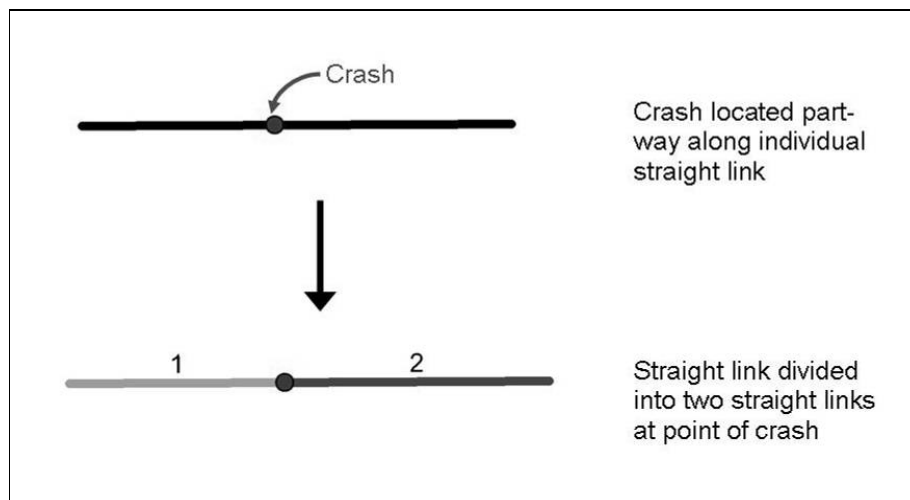
Firstly all arcs of the entire road network were divided into individual links, as illustrated for one arc in Figure 4.7. It should be noted that the entire road network, as opposed to only the roads within a specified buffer area around crashes, was processed so that calculations could later be applied to areas where crashes had not occurred.



**Figure 4.7 Example of Arc Division**

Links on which crashes occurred were also split at the point of crash, so that each crash point became a vertex in the road network, as shown in Figure 4.8. The majority of these “crash-links” were split using Lundeen’s (2005) program, although

links with multiple crashes required manual splitting as the program could not cope with such cases.



**Figure 4.8 Example of Link Division at Crash Location**

When arcs were split into individual links the attributes of the arcs were simply transferred to the individual links they created. This meant that some of the attributes, for example segment length, and identification number were incorrect. To counter this, a new ID field was added to the link table. Start and end vertices for each link were created and coordinates for each vertex were added to the vertices tables. These were then joined back to the link attributes table through the use of the new ID field. Thus each link had a unique ID and coordinates data for its start and end vertices, from which revised link lengths were computed. This processed data set is hereafter referred to as the “processed road network”.

In order to distinguish which vertices were also nodes (required for the weighting processes of some of the bendiness measures) the original road network, with polylines between nodes, was used. Nodes corresponding to the polylines on the network were created. A spatial join between the polylines and their nodes showed the number of polylines each node intersected. It was assumed that all other vertices from the processed road network joined two links. Thus the information between the nodes and vertices tables was joined through the use of identifying fields created from the coordinates of the points.

By this method, all links in the processed road network had information on the number of links (including the link itself) connected to each of its two vertices, termed “Start\_Int” or “End\_Int”. Cul-de-sacs and dead ends had a Start\_Int or End\_Int value of 1, intersections had values of three or greater and the majority of links had values of 2. A portion of the processed road network, along with associated attribute tables for the vertices and road links, is presented in Figure 4.9:



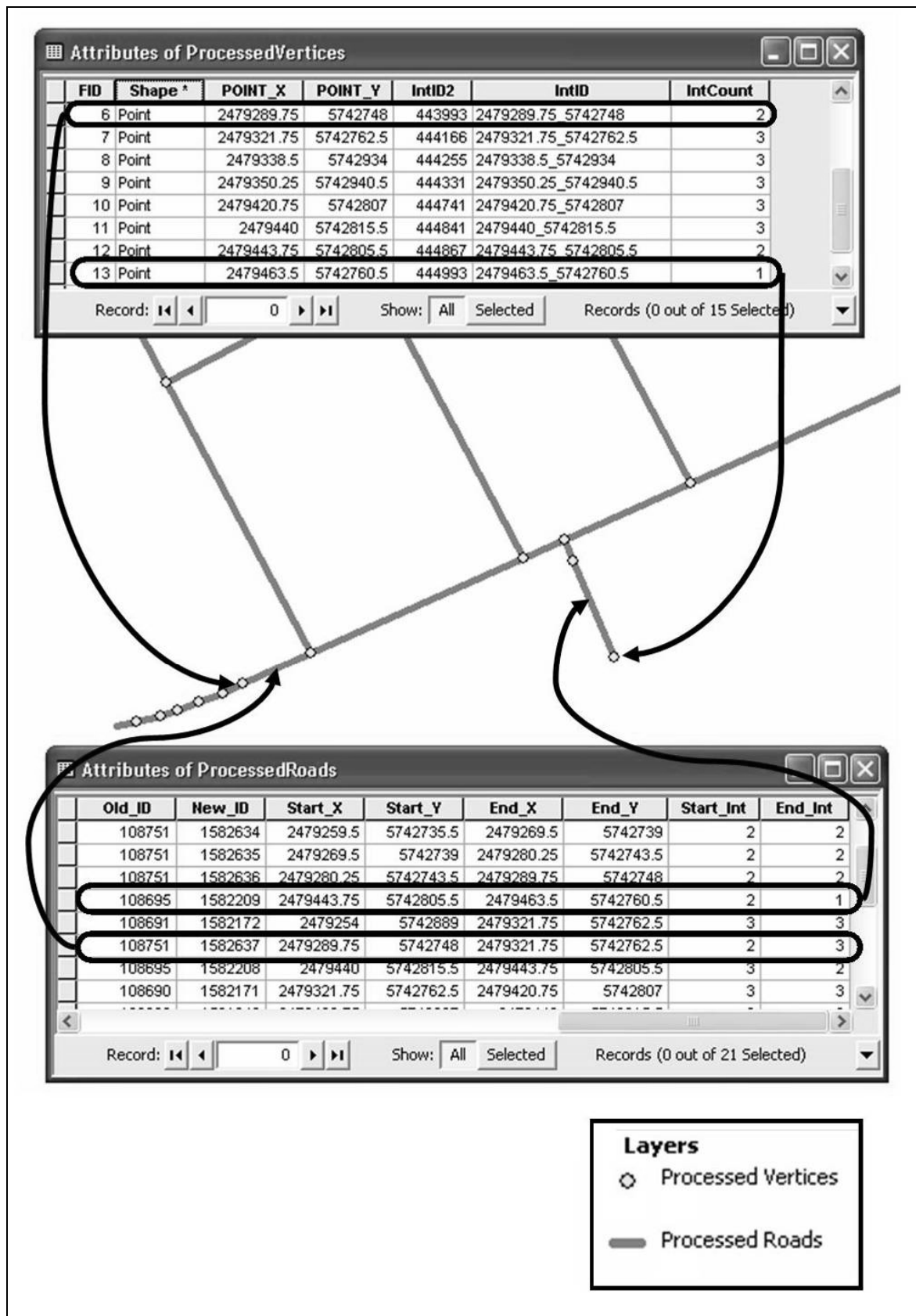


Figure 4.9 Part of the Processed Network

It was discovered, during the first stages of development of a bendiness calculation program using the processed road network that the original data contained some duplicate links. These links were not complete polylines but occurred where several polylines joined with a small overlap. Therefore it would not have been possible to remove duplicate links before the polylines had been split into straight links.

One occurrence of duplicate links was at Christchurch's Tunnel Road junction, as shown below in Figure 4.10, where selecting one section of the road with the pointer resulted in three polylines being highlighted.

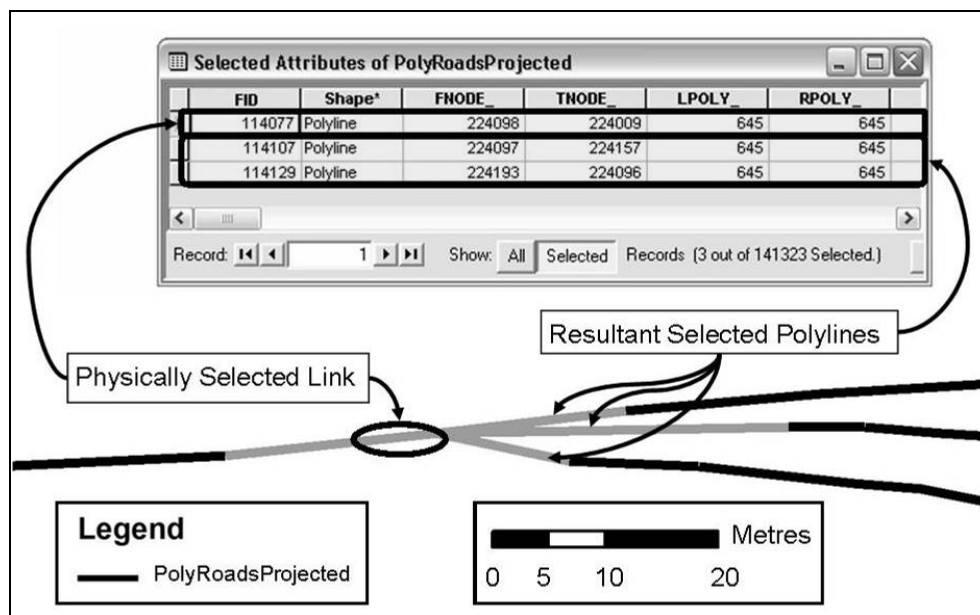


Figure 4.10 Duplicate Link in Original Data

This was undesirable as programs designed to progress along successive links of road would generally try to include all duplicate links, resulting in calculation errors (the program would effectively predict a driver travelling back and forth a stretch of road) or program failures. Thus, it was decided at this point to remove all duplicate links in the road network.

Several different programs were available to identify duplicates but, due to the size of the network, none were time efficient if used on the entire processed road network. It

was decided that the data should be subdivided according to one of its fields, such as TLA area or suburb names and duplicate links located in each individual section. One limitation of this method was that if two overlapping polygons each belonged to a different subdivision the duplicate link created would not be identified. Thus a balance was required between the time costs of using large sections and the accuracy costs of using very small sections.

As the duplicate links were created by overlapping polylines they occurred near to the nodes from the polyline data. Thus, a program was written to find, for each TLA, all the links within four links from a node (i.e. the links touching the node, the links touching those links, the links touching those links and the links touching those links). This selection was made into a new layer and reduced the size of the data to be checked for duplicate links significantly. This new layer was then processed with code based on that of the Siddal's (2005) code, which sorted through the records in the layer and, if more than one had the same geometry and coordinates, assigned a value of 1 to the specified "duplicate" field of all but one of those records.

All TLA files were then merged together and all records with a duplicate value of 1 were identified and deleted so that the road network then contained no duplicate links.

The vertices touching the duplicate links were then selected and a new spatial join performed to give new counts of the links touching the vertices where duplicate links had been removed. These corrected count data were then used to amend the Start\_Int and End\_Int fields of the processed network data.

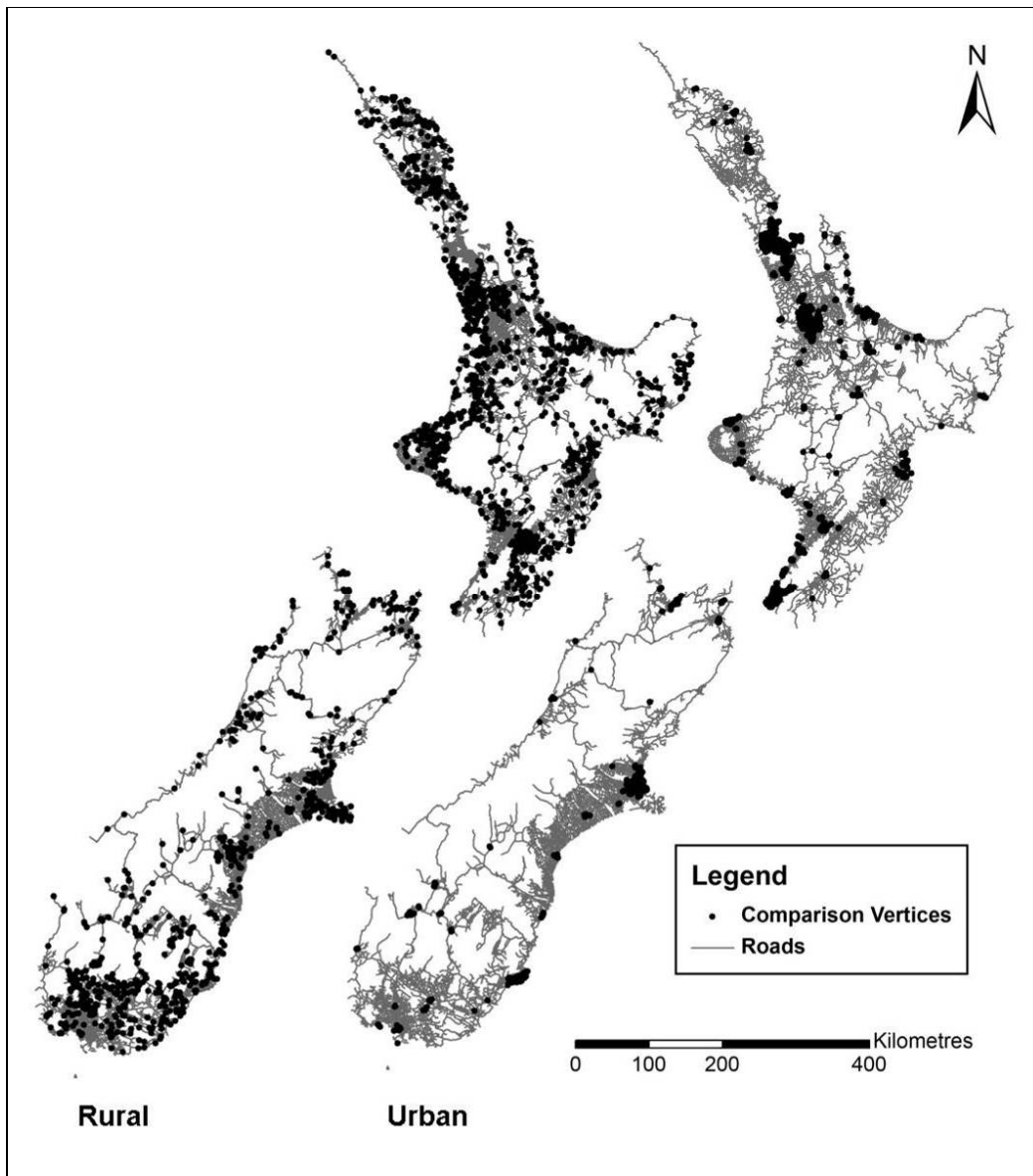
A network dataset was then created using the processed road network, with the new ID specified as the road name in the driving directions options. This meant that the each link on the network would be uniquely identified.

#### **4.3 Comparison Locations**

As discussed in Section 3.3, the statistical evaluation used for this study differed from that of the motivating study due to the difference in methodology. Results from the

analysis of areas around individual crashes were compared to characteristics of areas around non-crash locations to determine if certain levels of bendiness are more likely to contribute to crash occurrences.

The definition of a “non-crash location” required some thought. It was considered inappropriate to simply term all vertices that did not coincide with a crash to be non-crash locations as many of these vertices would still be very close to a crash. Thus, in keeping with the assumption that the length of influence would be one kilometre from the crash location, only vertices further than one kilometre from a crash location were considered. Of these vertices 2,000 rural and 2,000 urban vertices were randomly selected and used as substitutes for the crash locations in the bendiness measure calculations. These are shown in Figure 4.11.



**Figure 4.11 Rural and Urban Comparison Vertices**

After most of the bendiness calculations had been performed (as will be detailed in Section 5) the 2000 rural comparison vertices were sub-divided according to whether or not they were located on state highways in order to reflect the variation of flows found on rural roads. This was not the initial intention but as no flow data were available for the non-state highway roads it was deemed most appropriate to separate the two cases. This resulted in sub-sets of 219 and 1781 state highway and non-state highway comparison vertices respectively. While it is acknowledged that it would

have been more appropriate that these quantities match those of the crashes the comparison vertices were not re-selected due to the amount of time and computing power required for the calculation of the bendiness measures.

#### **4.4 Data Relating to Possible Other Influencing Variables**

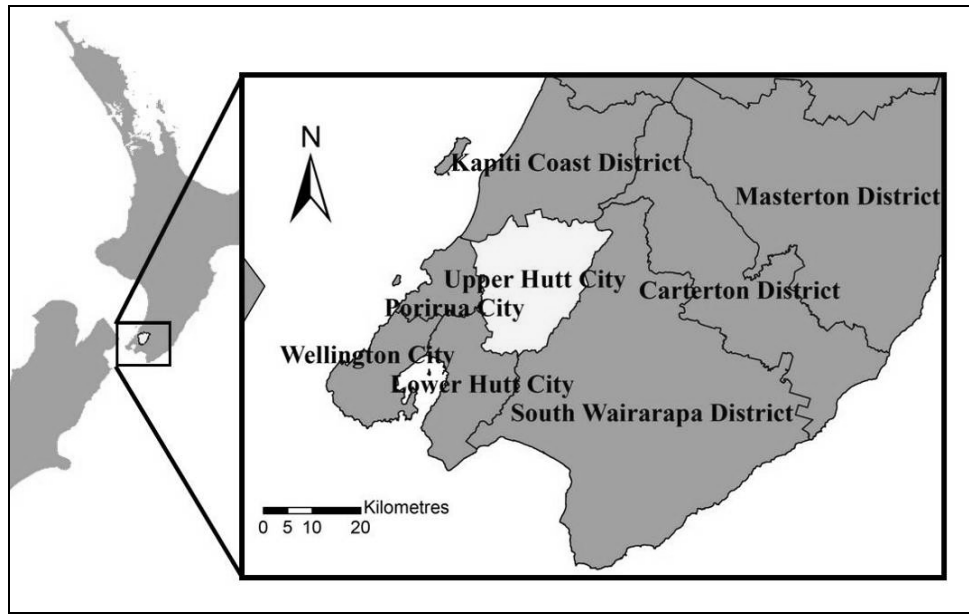
As for the motivating studies, it was expected that other influencing variables should be identified and accounted for in the statistical analysis. Variables tested are shown in Table 4.3 and were related to characteristics of the road network (either directly or as proxy measures) rather than socio-economic measures.

**Table 4.3 Data Sets and Sources of Possible Influencing Variables**

<b>Data</b>	<b>Description</b>	<b>Reference</b>
Upper Hutt City Flow Data	Upper Hutt City 2001 flow model	MWH New Zealand Ltd. (2001)
State Highway Flow Data	Average Annual Daily Traffic Count Estimations based on data from years 2002-2006	Transit New Zealand (2006)
Digital Elevation Model	Height above sea level for the entire country	Landcare Research (2000)
Rainfall Data	Mean monthly rainfall levels for entire country	Leathwick <i>et al</i> (2002)
Meshblock Population	Number of people who reside within each Census Meshblock area according to the 2001 New Zealand Census.	Statistics New Zealand (2002)
Meshblock Employment	Number of people who work within each Census Meshblock area according to the 2001 New Zealand Census.	Statistics New Zealand (2002)

#### 4.4.1 Upper Hutt City Flow Data

During the initial stages of research proposal, it was anticipated that information on the flows on each link of the road network would be required and that this may be obtainable from NZ CAS data. Further investigation revealed that comprehensive flow data for the whole of New Zealand would not be easily (or cheaply) attainable. Therefore a sample of flow data was obtained for the Upper Hutt City region and used as an example of what could be done if flow data for the whole country were available. Figure 4.12 shows the location of the Upper Hutt City TLA.



**Figure 4.12 Upper Hutt City Location**

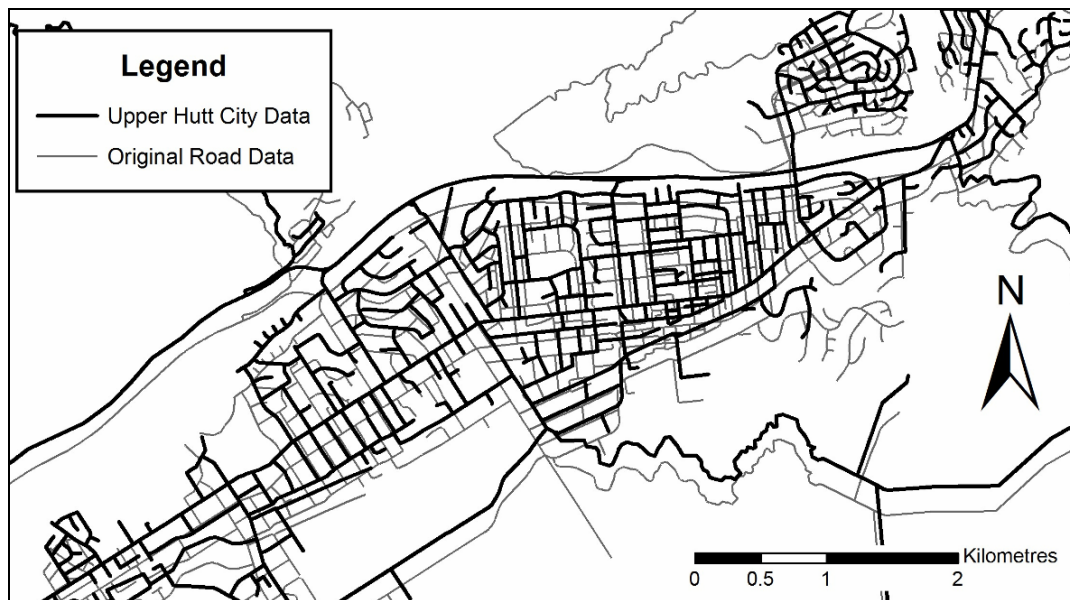
The flow data came as part of a road network projected with the “NZGD 2000 New Zealand Transverse Mercator” coordinate system. The network consisted of 2708 polylines with information on each link’s name, type and number of lanes. The flow data came from a model estimated from the 1996 New Zealand Census and calibrated with data from the 2001 New Zealand Census, along with information on average delays and speeds for each link.

The Upper Hutt network differed from the original New Zealand road network in several ways. Firstly, as mentioned previously the New Zealand road network was projected with the 1949 New Zealand Map Grid coordinate system whereas the Upper Hutt network used a different system. Thus there was an average distance of around 200m between corresponding points on the two networks.

Secondly, the two networks had been digitised independently so vertices occurred at different places along polylines and, disregarding the effects of using different projection systems, links were not matched exactly. Thirdly, some links existed in the Upper Hutt City network that were not present in the New Zealand network and vice-versa; this possibly reflected their different source dates.



These differences are apparent in Figure 4.13 which shows a sample of the two networks.



**Figure 4.13 Comparison of New Zealand and Upper Hutt Road Networks**

The final difference between the two networks was that the Upper Hutt network had a separate link for each direction of travel. Thus, for most links on the New Zealand Network there were two corresponding links on the Upper Hutt network, these links were geometrically identical but had different flow properties.

The flow data were recreated with combined directional flows and hence only one digital link for any given section of the physical road. These new data were then processed in a similar way to the original road network data; polylines were split into individual link and the midpoints of these individual links determined. A spatial join was then used to find, for each link on the original network, the characteristics (i.e. flow volume) of the closest flow network midpoint. Links that were not common to both data sets were discarded.

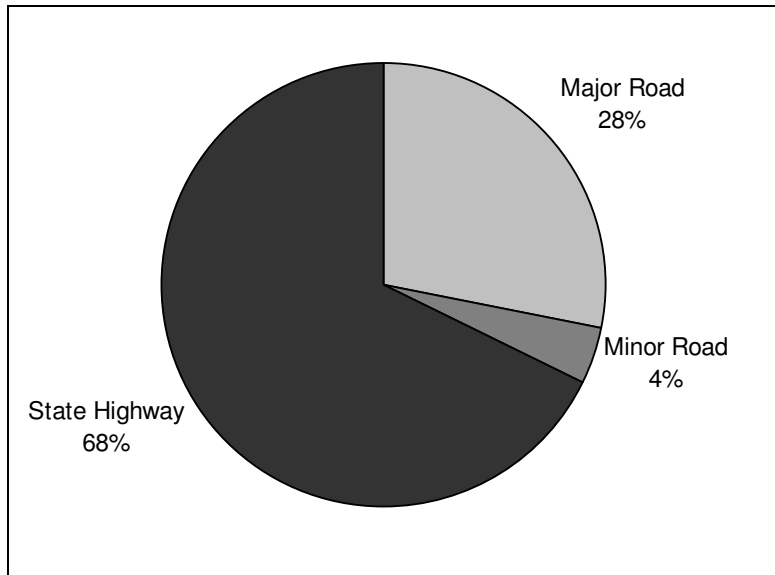
Thus the Upper Hutt City section of the original road network then included information regarding the flow on each link.

#### 4.4.2 State Highway Flow Data

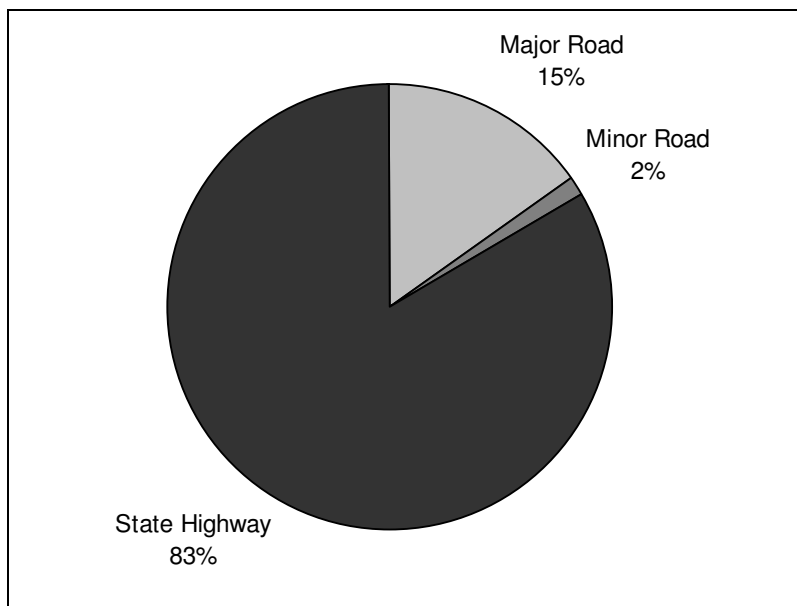
Transit New Zealand, the road controlling authority for the country's state highway network provided estimates of the average annual daily traffic flow on each section of state highway. The data were based on counts taken yearly from 1730 count sites for the period 2002-2006 inclusive.

As previously stated, no data were available for road types other than state highway on a national level. Generally state highways have larger volumes than other rural roads and are controlled by Transit New Zealand rather than a local authority. It was considered unwise to combine the rural state highway data with other rural road data as this distinction was the only one available to differentiate between sets of roads with different volumes. Also, the state highway flow data would be of no use if all rural crashes were combined but only some of their influencing roads had associated flow data.

Obviously, the influence areas of each crash occurring on a rural state highway may still contain non-state highway roads. However, the proportions of non-state highway roads within 1km influence areas of both the crash and comparison locations (which will be explained later on) are small, as shown by Figures 4.14 and 4.15. Most of the other roads are major roads and, as they are close to state highways, it would be reasonable to assume that they at least have similar flows.



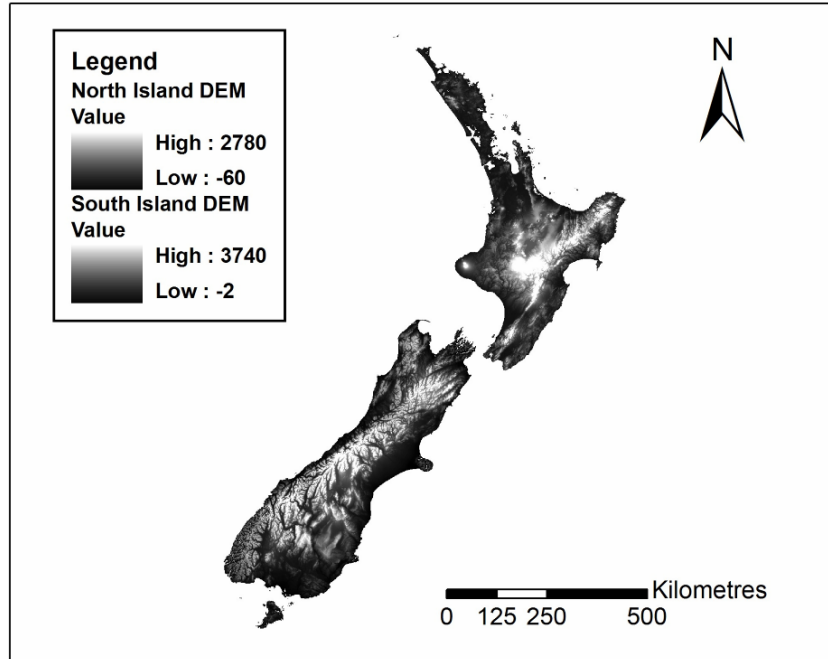
**Figure 4.14 Road Types for Rural Crash 1km Influence Area**



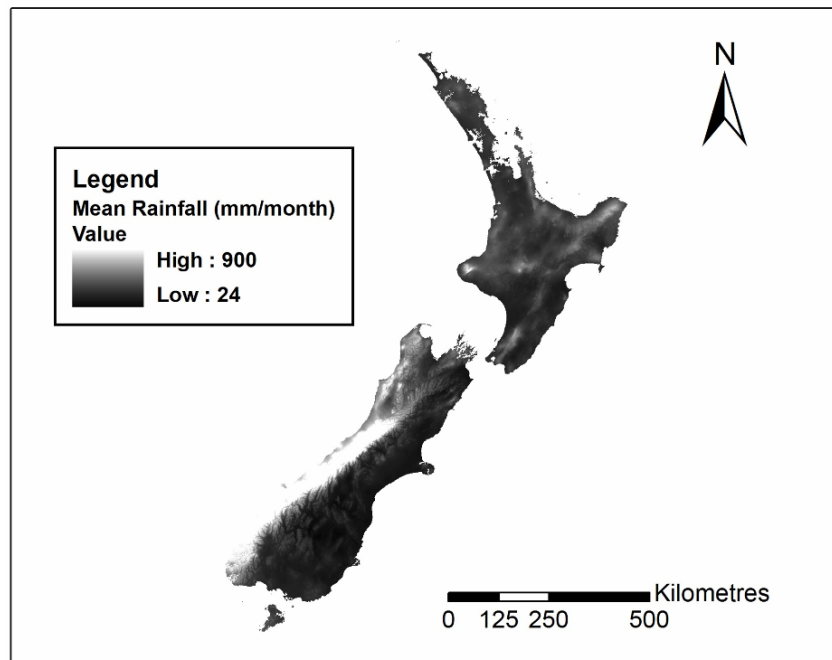
**Figure 4.15 Road Types for Rural Comparison Location 1km Influence Area**

#### 4.4.3 Elevation and Rainfall Data

Data to represent the elevation (Figure 4.16) and rainfall (Figure 4.17) within each crash's area of influence came from the same sources as that used in the motivating study.



**Figure 4.16 North and South Island Digital Elevation Models**



**Figure 4.17 Mean Precipitation Model for New Zealand**

The values of elevation and mean precipitation for each vertex that fell within the area of influence of each crash and comparison vertex were collected and used to find the mean and standard deviation of elevation and precipitation corresponding to each

location. Ideally the values would have been weighted according to each vertex's tributary length, but as this would have taken a lot more computing time and power and within a 1km route link lengths were generally consistent it was decided that the faster method would be reasonable.

#### 4.4.4 Employment and Population Data

In order to account for the differences in flow in different areas where flow data were not available it was assumed that other characteristics could be used as proxy measures. Trip generation models are generally based on the number of people who live and work within an area; the higher these values are the higher flows on surrounding roads are likely to be, particularly in urban locations. It was expected that these measures would not be as useful in rural situations where traffic is often passing through rather than starting or ending a trip.

Information regarding the number of people who live and work within each of the country's census meshblocks was obtained from New Zealand's 2001 Census (Statistics New Zealand 2002). Each crash and location site was then assigned the population and employment properties of their meshblock through a spatial join operation. Meshblocks boundaries are assigned in a way that ensures each contains a similar population (roughly 60-100 residents) so it was necessary to weight the population and employment values for each meshblock by its area to give an indication of the respective densities of population and employment.

#### 4.4.5 Junction Density

It was considered that the number of junctions within a location's influence area may affect the crash occurrence. Two junction density measures were calculated for the 1km influence area surrounding each crash and comparison location. A junction was defined as any vertex with three or more adjoining road links. Hence all junctions were nodes, but some nodes (for example cul-de-sacs) were not junctions. The first method simply divided the number of junctions within the influence area by the total length of roads within the influence area:

$$JD = \frac{J}{d}$$

**Equation 21 Junction Density Calculation**

Where:  $JD$  = the Junction Density value for a particular location

$J$  = the number of junctions within the influence area

$d$  = the total length of links within the influence area

The second method took into account the fact that junctions with many legs are likely to be more hazardous (or induce more caution) than more simple junctions:

$$\overline{JD} = \frac{\sum_{j=1}^J l_j}{J \times d}$$

**Equation 22 Weighted Junction Density Equation**

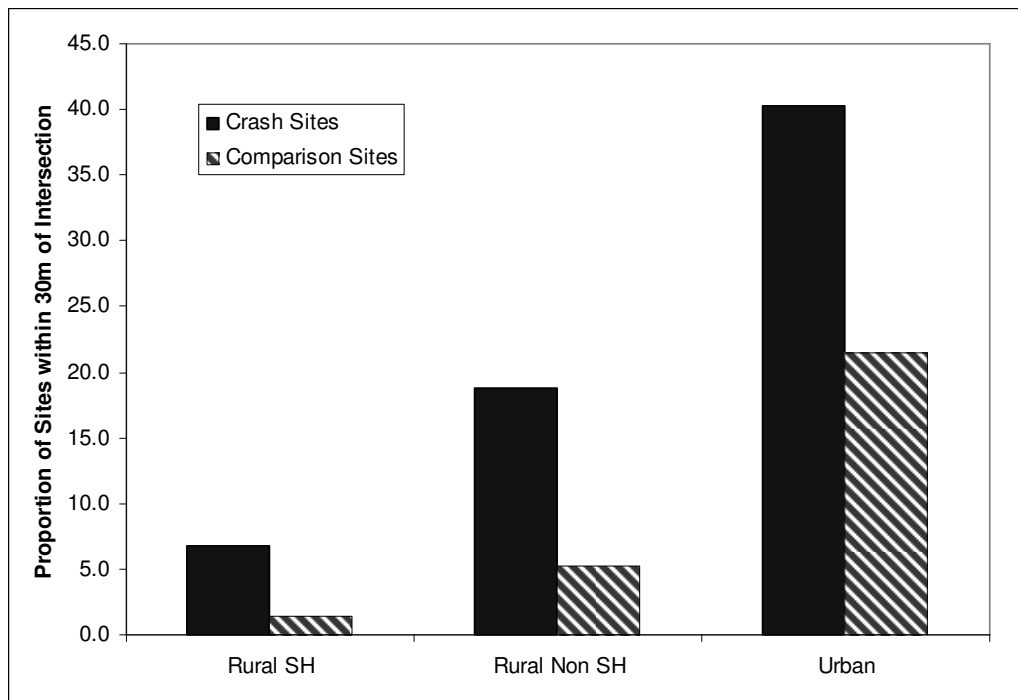
Where:  $\overline{JD}$  = the weighted Junction Density value for a particular location

$l_j$  = the number of legs at junction  $j$

4.4.6 Intersection versus Mid-block Locations

It was assumed that whether or not each crash (or comparison site) was located at an intersection would have an effect on the model, given that intersection and non-intersection crashes generally involve different influencing factors. It was identified previously that the CAS intersection classifications were not always consistent. Thus a GIS search was used to find all crash and comparison locations within 30m of an intersection, according to the LTNZ (2004) definition of a crash site. The Ministry of Transport's (2007) definition of an intersection crash being within 10m of an intersection (as used in CAS) was not used. It was speculated that, due to the influence of intersections on the surrounding roads and inaccuracies of the spatial data, a large area would be more appropriate.

It can be seen in Figure 4.18 that the urban sites had a much greater proportion of intersection crashes than the rural sites, especially the state highways, which would be expected given the increased junction densities in the urban situation. However, the comparison sites in all three cases were much less likely to occur at intersections than the crash sites, which may indicate a flaw in the choice of comparison sites.



**Figure 4.18 Proportion of Intersection Sites in Crash and Comparison Samples**

#### 4.4.7 Variables Intentionally Excluded from Investigation

It was not within the scope of this research to account for effects of socio-economic characteristics (such as age, ethnicity and income measures) on crash occurrence. This was primarily because, from a traffic engineering viewpoint, variables that can be in some way controlled are more important than those that cannot. Bendiness itself is seen to be something that can be regulated through specifications for new road design and remedial actions to existing roads. Environmental factors likely to increase crash risk, such as high rainfalls that result in decreased skid resistance, can be combated by using specialised treatments in those areas. Socio-economic characteristics, however, cannot be controlled by traffic engineers.

Secondly, the availability of data was a limiting factor. The motivating study used socio-economic data at a TLA level of aggregation which was consistent with the method of calculating bendiness measures on a TLA basis. This study however focused on individual crash locations and would therefore require much more localised socio-economic data.

It would have been possible to obtain data at a census meshblock level of aggregation however this idea was rejected as it would imply that the socio-economic characteristics of residents surrounding crash locations were influential, as opposed to the socio-economic characteristics of the drivers actually involved in the crash. While the two groups may sometimes have coincided (as previously mentioned, a significant proportion of crashes do occur close to drivers' homes) there was no way of distinguishing when this occurred.

It was postulated that some of the socio-economic measures used in the motivating study may have been linked to more tangible road network characteristics. For example regions with lower mean incomes will have less money available to spend on road maintenance and may naturally be in areas with worse terrain. Based on correlations of the socio-economic and terrain measures at a TLA level used in the motivating study there is little evidence to support this theory, however a more in-depth investigation may prove otherwise.

#### 4.4.8 Desired but Unattained Data

Not all the types of data that were wanted were attainable. For example, it would have been appropriate to have flow data for the whole of the country's road network similar to the state highway flow data. As has been mentioned, the meshblock population, employment and area data were intended as a proxy for this.

Also, it would have been useful to have information regarding the routes taken by drivers previous to their crash. This would have removed the need to weight different route options and may have reduced the necessity of having comprehensive flow data.



## 5.0 Case Study Bendiness Measure Calculations

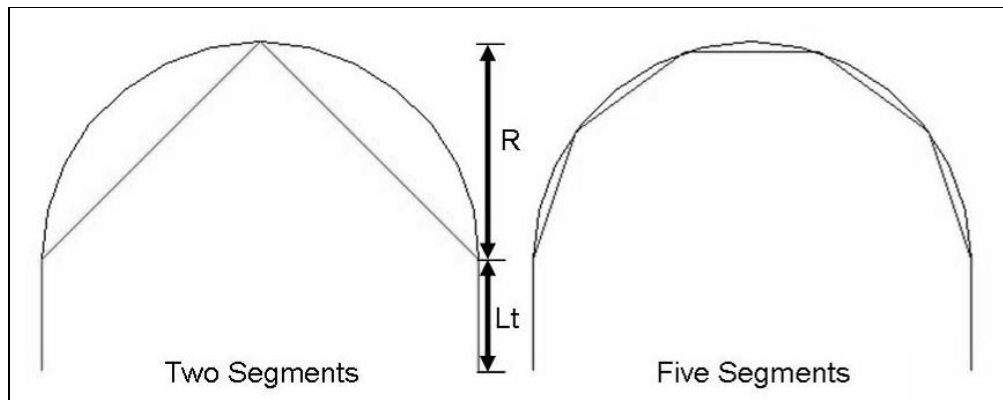
This section outlines the steps taken to calculate various bendiness measures for the New Zealand case study. Four of the measures used by Haynes *et al.* (2007a) in their British study and the motivating study (bend density, detour ratio, cumulative angle and mean angle) were modified to suit the case study and used as a starting point. From this, new bendiness measures were also created and tested. Flow data for a subsection of the case study road network were obtained and this information used to improve the bendiness measures calculations.

### 5.1 Applicability to Case Study of Bendiness Measures used in Motivating Study

It was initially assumed that the straightness index as used by Haynes *et al.* was not an appropriate measure. A preliminary investigation of the Canterbury region showed that only about 0.2% of the links passing within a 1km buffer zone of the crashes were greater than 1km in length (i.e. completely straight lengths). Hence 99.8% of the crashes would produce a straightness index of 0. Thus it was decided that the straightness index used in the motivating studies would not be a useful measure of bendiness for this application. Subsequently the team working on the motivating study reached the same conclusion and the straightness index was omitted from in their further analyses.

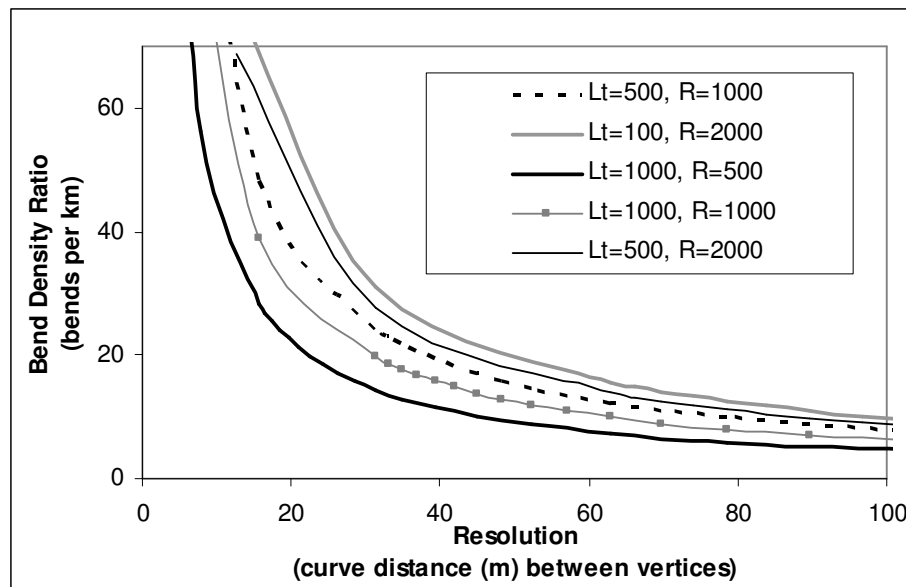
It was also assumed that values obtained from the bendiness measures would depend greatly on the resolution (spacing between vertices) of the digitised road network. It was expected that the resolution of the road network would not be constant throughout the data set, due to it coming from several different sources and the fact that digitisation is an imprecise process with accuracy depending heavily on the person doing the digitising.

To test the effects of road network resolution on the outcome of the bendiness measures, a simple method of approximating a portion of semi-circular road with adjoining tangents by a series of straight lines was implemented, as illustrated in Figure 5.1, where the curve of the actual road has been replaced firstly with a two-segment approximation and secondly with a five-segment approximation. It can be seen that the more segments used, the better the approximation but there will always be an error involved.



**Figure 5.1 Approximation of Circular Curve with Straight Sections**

The bendiness measures for different digitising resolutions and different curve geometries (the radius length, “R”, and tangent length, “Lt”) were derived by spreadsheet calculations. These are shown in Figures 5.2 to 5.5. It was concluded that length calculations were less affected by variations in scale than angle calculations and that the detour ratio was the bendiness measure least sensitive to changes in scale.



**Figure 5.2 Effect of Resolution on Bend Density Ratio**

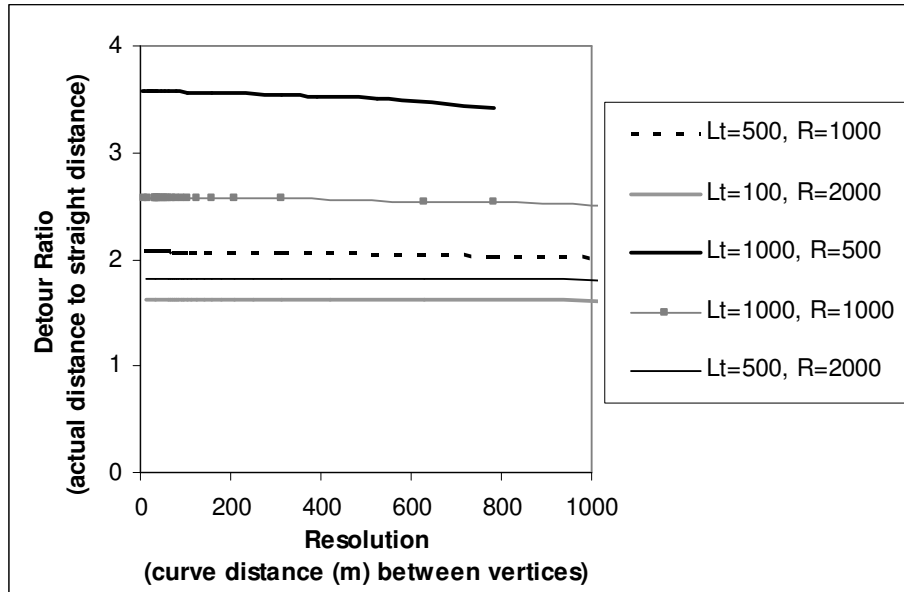


Figure 5.3 Effect of Resolution on Detour Ratio

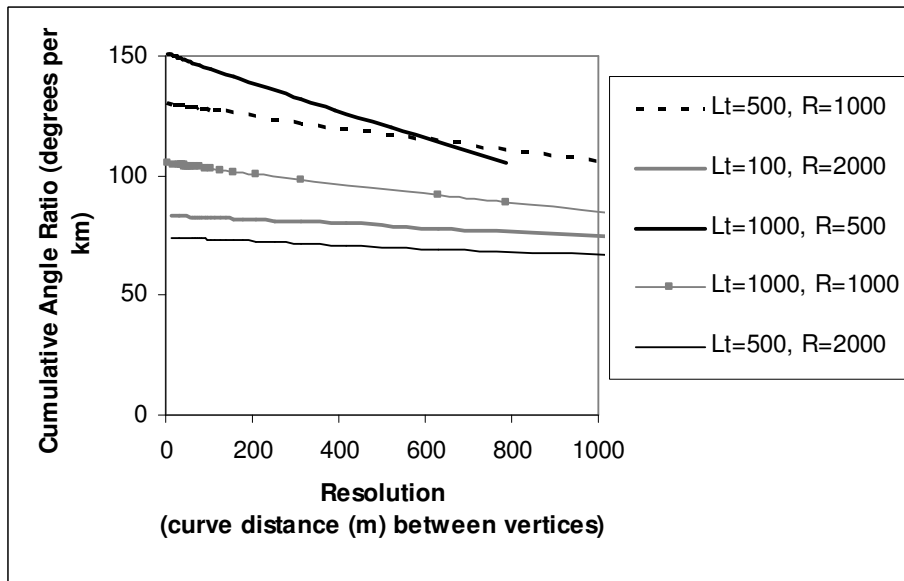
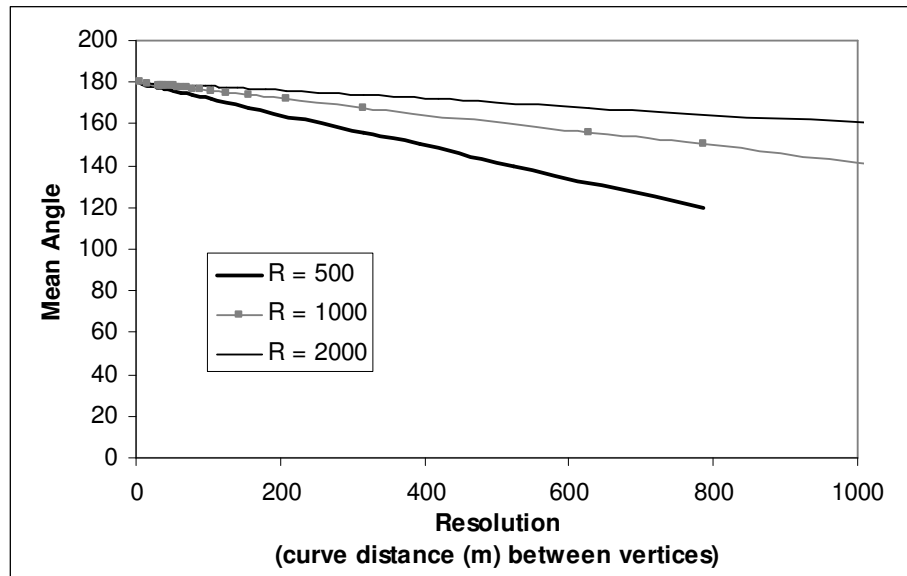


Figure 5.4 Effect of Resolution on Cumulative Angle Ratio



**Figure 5.5 Effect of Resolution on Mean Angle Ratio**

Two preliminary conclusions regarding the applicability of the bendiness measures as used in the motivating study were made based on the results of this theoretical section of road. Firstly, as the digitised road network combined small networks from several different sources it is important that the different sources have similar resolutions. Otherwise meaningful comparisons of different bendiness measures at different locations cannot be made. The cumulative angle and mean angle measures would also be appropriate within certain ranges of resolution, but the bend density ratio should not be used if resolution varies. The detour ratio was the most consistent over a range of resolutions so would be the most appropriate for this study where the amalgamated road network did not have one consistent resolution.

Secondly, bendiness values obtained from this study will not necessarily be transferable to actual road geometry or international comparisons. This is because the bendiness measures have been shown to vary significantly with the resolution of the digitised road network. It would not be sensible therefore to make specifications for actual road design without somehow also specifying a way of using resolution in the design. For example, consider the scenario that the study finds that stretches of road with cumulative angles of under 100 degrees per km have lower crash rates. It would be useful to pass this information on to designers and specify some way of ensuring new roads do not have cumulative angles of over 100 degrees per km. However, this

value would be heavily dependant on the resolution of the digitised road network analysed. In order to actually check the cumulative angle of a designed road it would have to be digitised to the same level as that used in the study and then conduct a cumulative angle evaluation.

The relative ordering of the different length-radius combinations, however, did not change with resolution. Hence, studies concerned with gauging the relative rankings of different site types should be more dependable than studies concerned with gauging absolute rankings.

Similarly, it may be desirable to use the results of the study to predict trouble spots that may be hazardous due to their bendiness properties. This would be easily done for New Zealand locations by using the same digitised road network used in the study. A simple GIS search could identify stretches of road with undesirable properties. However, overseas traffic engineers would need to ensure that the resolutions of their digitised road networks were similar to the one used in this study before using the same bendiness values. This could be overcome by finding relationships between the bendiness measures and resolution. However, the relationships in the theoretical analysis were solution specific; none of the bendiness measures versus resolution graphs had equal gradients between the different geometric situations.

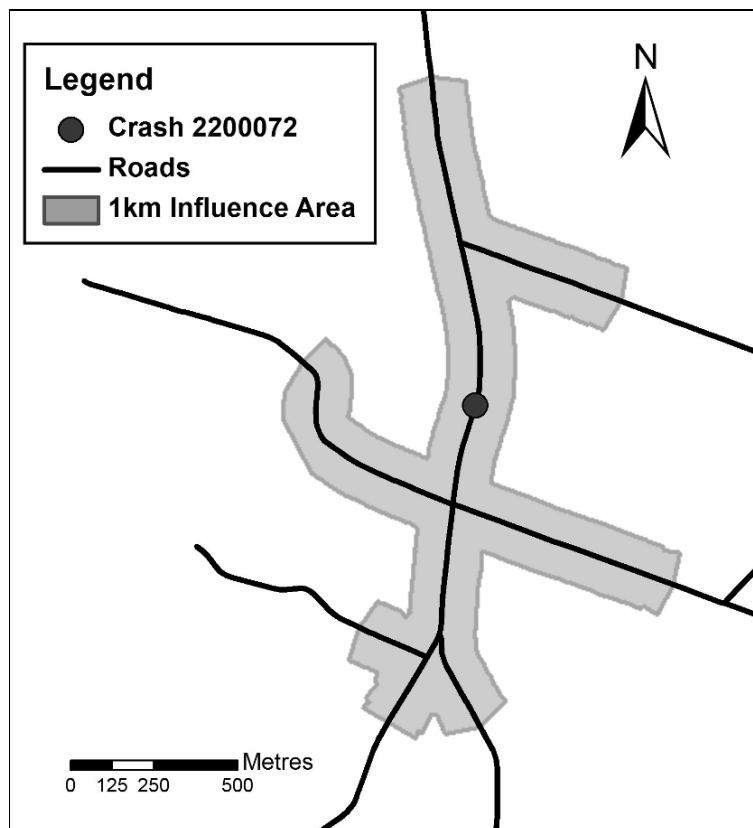
It was anticipated that, once the use of existing bendiness measures on a crash-by-crash basis had been implemented, alternative methods that take into account changes in speed environment of successive stretches of road would be investigated. One postulated way of doing this was to calculate the variation in change in angle between each successive vertex along a route. A high variation would indicate a low level of design consistency.

## **5.2 Calculation of Bendiness Measures used in the Motivating Study**

Calculation of the bend density, detour ratio, cumulative angle and mean angle was performed by creating Visual Basic for Applications (VBA) programs to be used in ArcMap (ESRI 2006). The straightness index was not calculated for reasons outlined previously.

The bendiness calculation program was designed to use the processed road network and the relocated crash data. The initial preliminary step was to create the additional fields required for calculations and output (this step could be skipped if the attribute tables already contained these fields).

The first crash in the crash table was selected and its coordinates determined. The next step required the selection of all road links in the influence area of the crash by creating a “service area” polygon to extend one kilometre around the crash. This dimension was chosen based the work of Koorey and Tate (1997) which used segments one kilometre long for evaluation of New Zealand roads. This is illustrated for the influence area of crash 2200072 (ID number from the CAS database coding) in Figure 5.6:



**Figure 5.6 1km Influence Area around Crash**

It was initially intended that the effect of using different sized influence areas would also be tested. It was thought that influence areas based on a set travel time rather than distance would be more appropriate. However, due to time and resource constraints, this was not investigated fully and has instead been recommended as a possibility for future research in Section 8.4.1.

The bendiness measures were to be calculated between the crash and each of the vertices immediately outside of the influence area or vertices within the influence area that connected to only one link (cul-de-sacs). Two assumptions were made here, firstly that when multiple route choices for getting from the crash location to an external node existed, the shortest route was taken. And secondly, that in most cases the driver was travelling a distance greater than the route length and hence originated from outside of the influence area.

Obviously these assumptions were not completely realistic. Drivers base their route choice on factors other than distance; it would have been more appropriate to consider travel time but this would have required data on traffic flows and speed limits and a much more comprehensive network analysis approach. Also, it is possible that drivers originated from within the influence area. An American study of 11,000 reported crashes showed that 23% of crashes occurred within 1mile (1.6km) of the driver's home (Progressive Insurance 2002), although this includes drivers returning home as well as originating from home. Even so, this means that an infinite number of possible origins existed within the influence area but were not considered. It was necessary to make these assumptions in order to proceed, due to the limitations of data and resource availability. Cul-de-sacs within the influence area were included to reduce the likelihood of incorrect weightings being applied to the routes.

To find the external vertices (those immediately outside the influence area) all the links that at least partially intersected the influence area were selected and from this selection all links completely contained within the influence area were removed. All vertices intersecting the remaining links were selected and of those, vertices contained within the influence area were removed from the selection. This selection was saved as the external vertex layer corresponding to the particular crash and was added to the map, as illustrated in Figure 5.7.

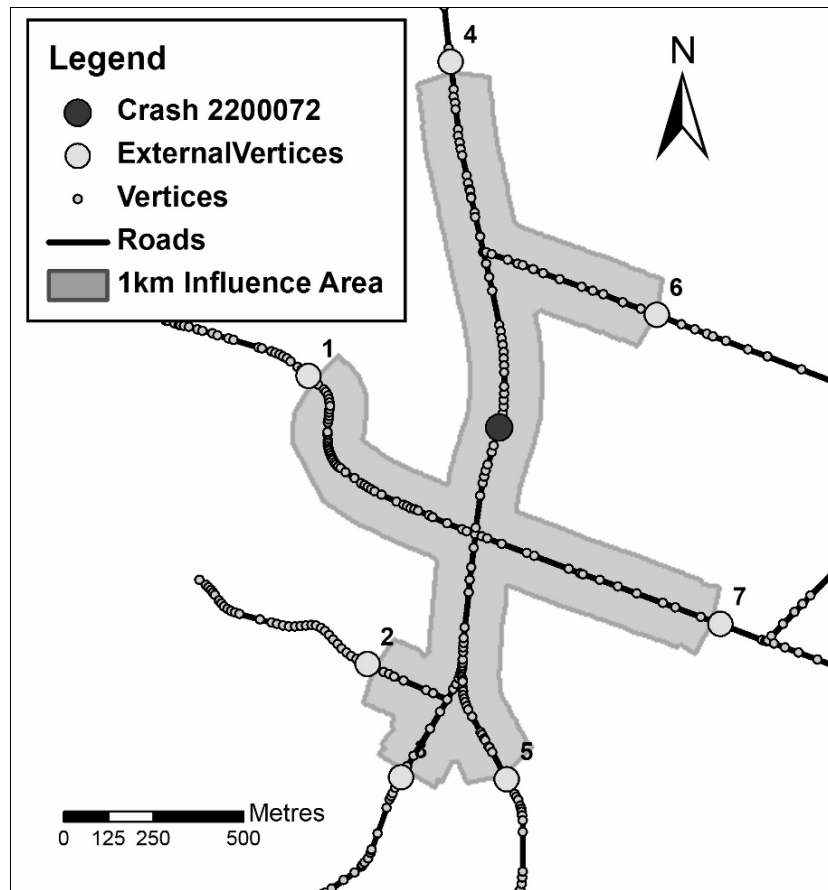


Figure 5.7 Identification of External Vertices

Next, a route layer was created between the crash and the first vertex in the external vertex layer, as shown in Figure 5.8.



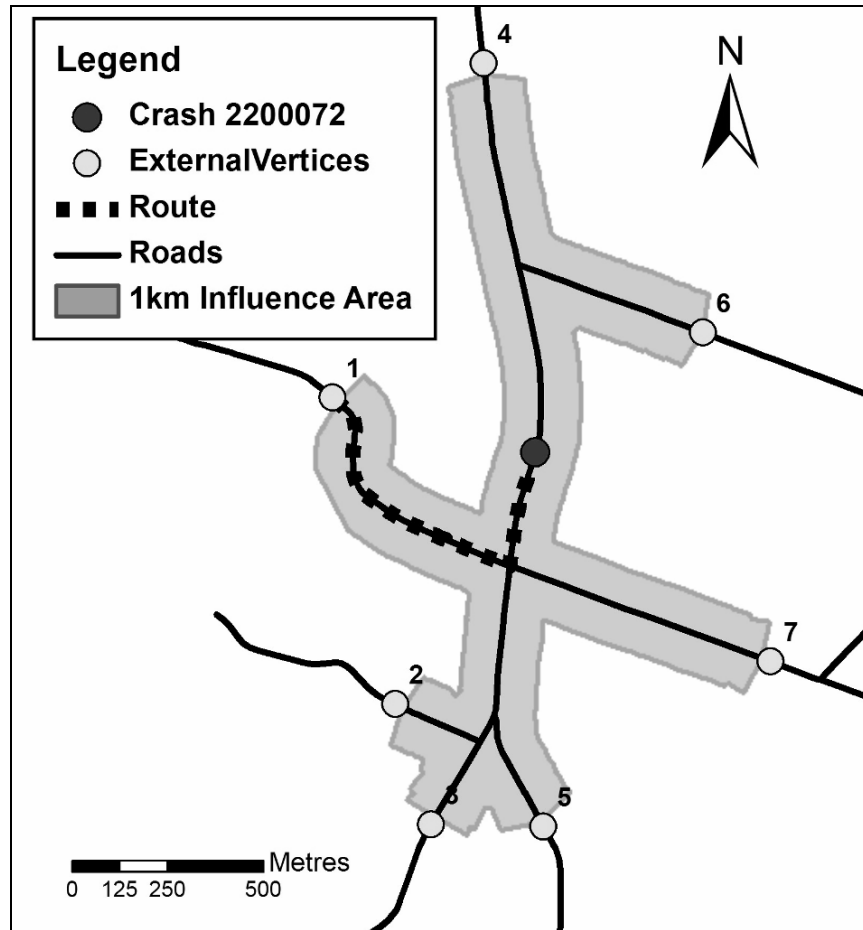
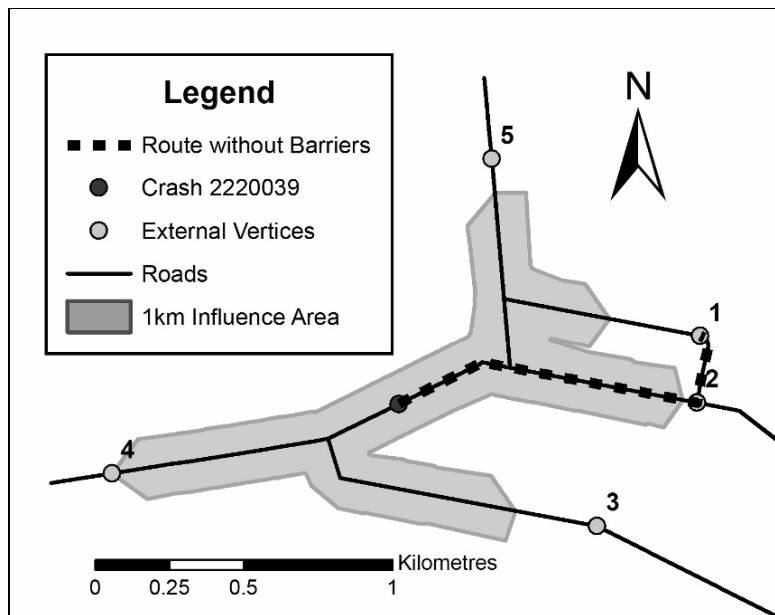
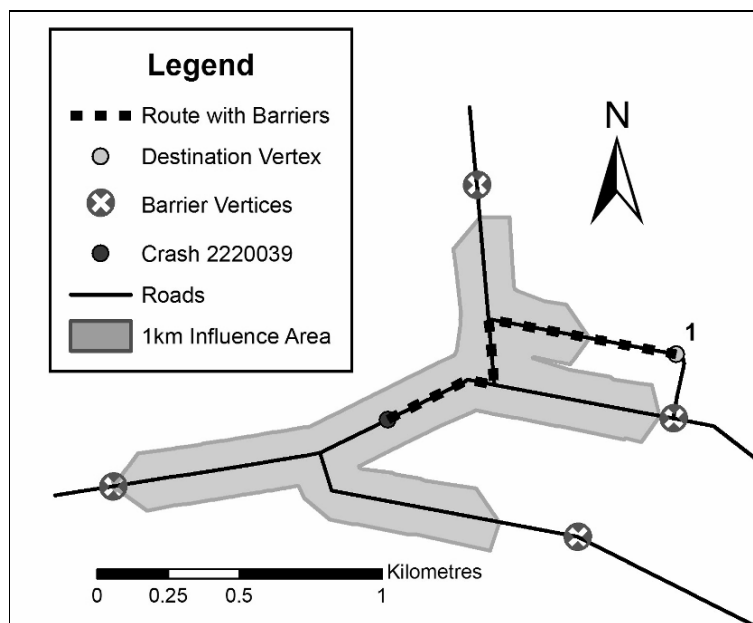


Figure 5.8 Example Route

It became apparent that, in some situations, where the external vertex was a long distance from the edge of the influence area it was possible for a shorter route than the one intended to be created between the crash and the external vertex, thus incorrect bendiness measures would be obtained. The incorrect route would have to pass by one of the other external vertices in order to reach its destination. Thus, to avoid this problem, all other external vertices were added as barriers in the route layer, so the only route that could be created would be the desired one. This is illustrated in Figures 5.9 and 5.10 for the route between crash 2220039 and the first external vertex. The route in Figure 5.9 is obviously wrong as to get to external vertex 1 it passes through external vertex 2. The route in Figure 5.10, achieved by placing barriers at all external vertices except the target vertex one, is the correct route.



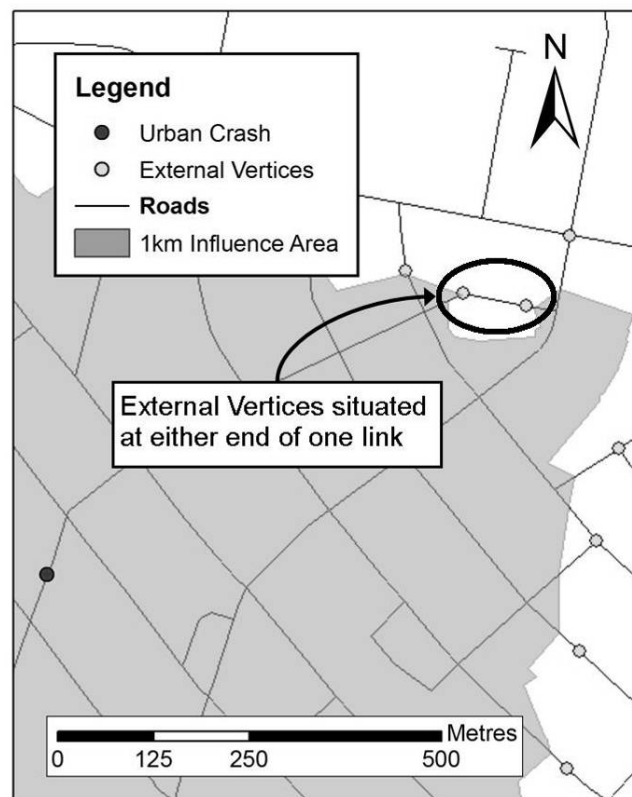
**Figure 5.9 Incorrect Route without using Barriers**



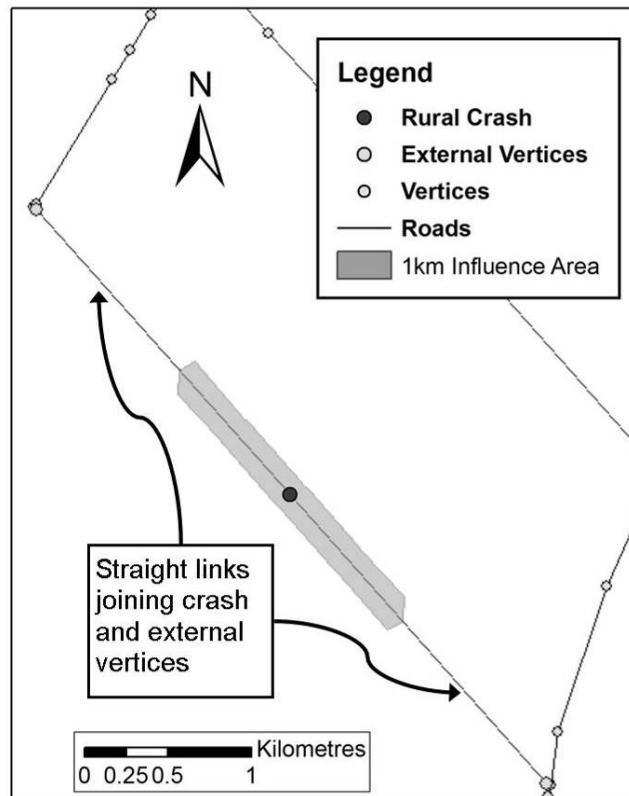
**Figure 5.10 Correct Route using Barriers**

Unfortunately, the method of using barriers introduced new errors. ArcMap does not assign the non-traversable properties of a barrier to the corresponding vertex but to one of the links it touches. (There was no practical way of choosing which link would be made non-traversable when several options existed). Thus if a location was

situated on a link that had a barrier at the other end it was possible that the link would be non-traversable and the location would become unreachable. This occurred mainly in urban situations, where external vertices (and hence barrier-location combinations) were very close together due to small network links, for example the link indicated in Figure 5.11. A similar problem also occurred occasionally in rural situations where the road links surrounding a crash were longer than the route distance and hence the external vertex and crash location were situated on the same link, an example of which is shown in Figure 5.12. It was decided that all sites would be processed using barriers and any for which the method was not effective would be re-processed without barriers.



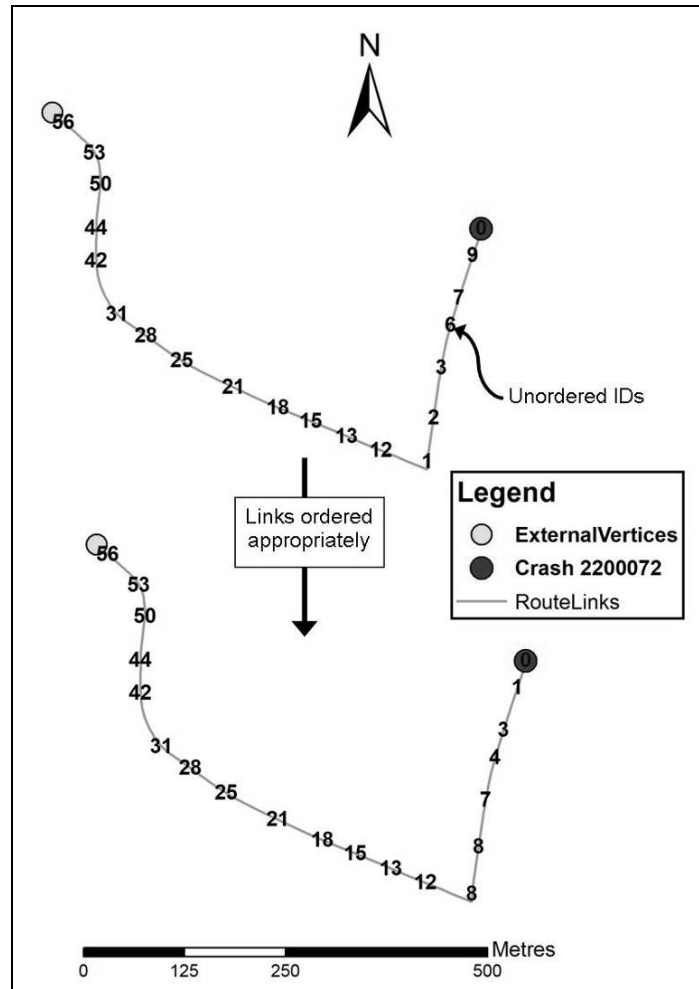
**Figure 5.11 Urban Situation where Barriers are Unwanted**



**Figure 5.12 Rural Situation where Barriers are Unwanted**

Once the route had been created, all links from the road network that had their centre within 0.1m of the route were selected. The 0.1m tolerance was required because without it very short links on tight curves were often excluded.

The selection was saved to a new file named the “route links” file and was added to the map. The links selected were not necessarily in order so the table had to be rearranged so that the bendiness measures could be calculated using successive links. The route links table was sorted through to find the first link (i.e. that with coordinates at one equal to the crash coordinates) and from that successive links. Each link was assigned an “order” integer value corresponding to its position in the route, this is illustrated for the first route of crash 2200072 in Figure 5.13.



**Figure 5.13 Ordered Route**

ArcMap automatically defines the object IDs and these cannot be modified by users. Thus, to store the route links in an appropriate order, a new ordered route links file with the same fields as the original route links file was created and records were filled so that the order number corresponded to the object ID number in the new table. Once this was done, the bendiness measures for the route were calculated between successive links and stored for the whole route.

The route steps were repeated for each of the remaining external vertices in the crash's external vertex table. Once all routes had been created their bendiness measures were aggregated to give total bendiness measures for the crash. The weighting method described in Section 3.2.3, whereby each route was weighted according to the number of links at each of its intersections was used for all the

models. It was not possible to use the flow-weighted method describe in Section 3.2.4 even for the rural state highway crash model since, as shown in Section 4.4.2, not all of the roads in the influence areas of the rural state highway crashes were state highways.

All steps were then repeated for the next crash in the crash data and so on until bendiness measures for all routes and hence for all crashes had been computed.

### **5.3 Additional Bendiness Measures created for Case Study**

It was assumed that the bendiness measures used previously stood as proxy measures of design consistency as they evaluated the range of geometric deviation of a network. From the previous studies it could be inferred that regions with extremely high or extremely low values of bendiness have high levels of design consistency (i.e. they are either consistently bendy or consistently straight). However, it was not clear if regions with medium levels of bendiness were comprised of smooth, consistent curves or combinations of sharp curves and long straights. Hence additional measures to account for this were attempted.

#### **5.3.1 Standard Deviation between Angles**

In order to give an indication of the variation of elements along a route the standard deviation of angles along a route was calculated. This was performed in the case study at the same time as the bendiness measures adapted from the Hayne's studies, as detailed in the previous section. The standard deviation of angles was calculated according to the formula:

$$SD = \sqrt{\frac{\sum (\theta - MA)^2}{(n - 1)}}$$

**Equation 23 Standard Deviation of Angles**

where:  $\theta$  = an angle between two links

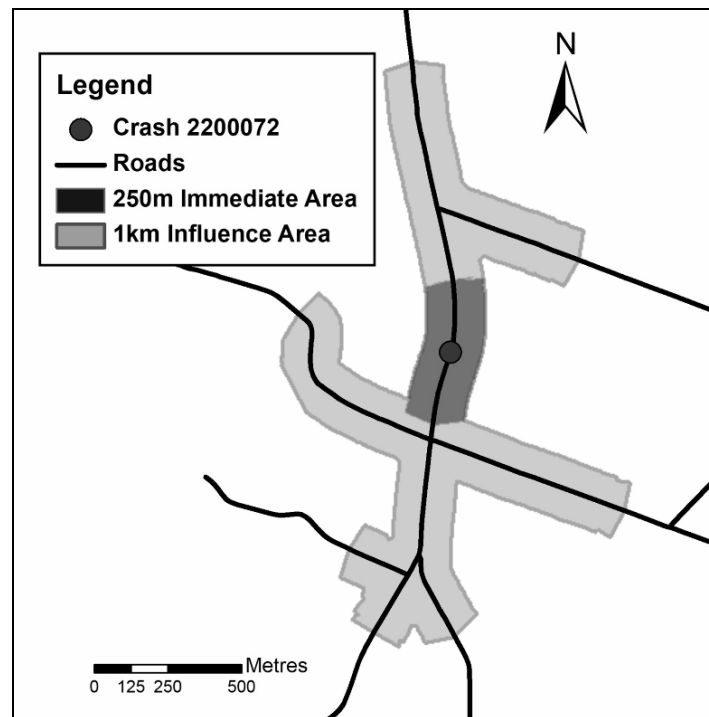
$MA$  = the mean angle for the route

$n$  = the number of angles the a route

This measure, while often being classed as another bendiness measure for the purposes of this report, was most importantly a measure of the consistency of angles within a section of road. Thus it was expected that the standard deviation of angles terms might have a different relationship to crash occurrence than the bendiness measures used by the motivating study.

### 5.3.2 Comparison of Measures for Immediate Area versus Influence Area

Another method of considering design consistency was attempted, based on the theory that sections of road that are very inconsistent with preceding sections are more likely to induce crashes. This involved examining the bendiness characteristics of the crash or comparison point's immediate vicinity with the bendiness characteristics of the 1km influence area. For this exercise the immediate area was defined as the parts of the network within 250m of the location of interest, as illustrated for crash 2200072 in Figure 5.14:



**Figure 5.14 Immediate and Influence Areas for Crash 2200072**

Ratios for immediate versus influence bend density, detour ratio, cumulative angle, mean angle and standard deviation of angles were obtained for rural crash and

comparison sites. This effectively normalised the bendiness measures for a particular site; a ratio of one signified that the bendiness of the immediate area was the same as the that of the influence area. Because each site was expected to have similar flows and other non-bendiness characteristics for the immediate and influence areas there was no need to consider the normalised bendiness measures in conjunction with other influencing variables.



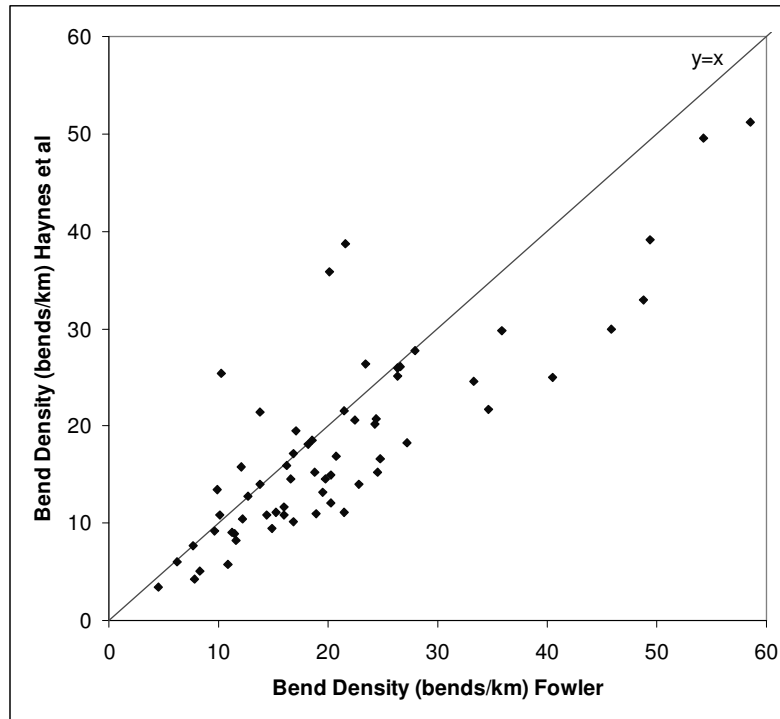
## **6.0 Preliminary Case Study Investigations**

This section compares the method of this research (termed here the “Fowler method”) with the method of the motivating study (the “Haynes method”) at a TLA level for both rural and urban crashes. Next the Fowler bendiness measure results are compared for crash and non-crash locations. Finally, the limitations of assessing the bendiness of sites without using flow weighting for different routes are investigated by comparing two weighting methods for the Upper Hutt City TLA. These three investigations were performed to assess the viability of the Fowler method and provide information for subsequent research.

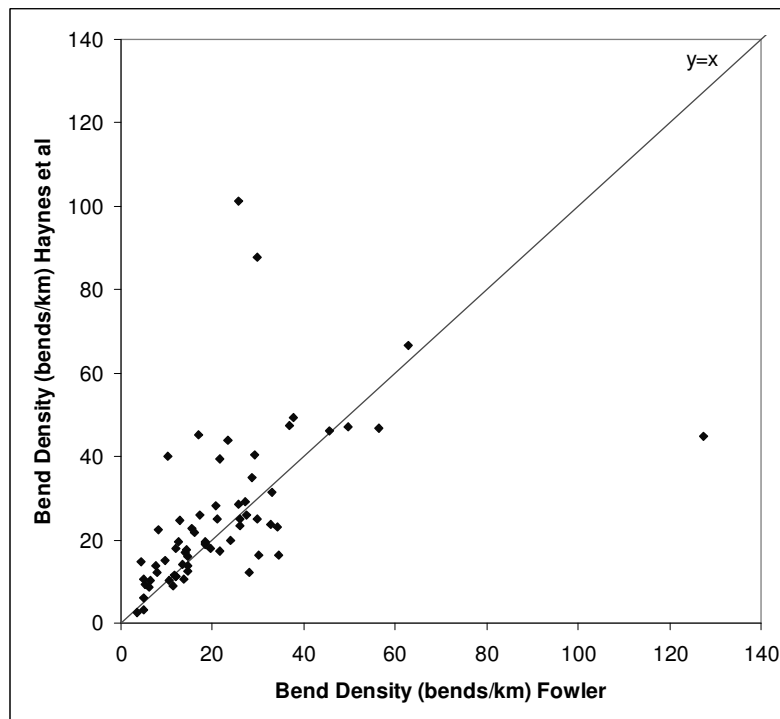
### **6.1 Comparisons with Motivating Study**

#### **6.1.1 Rural Road Crashes**

Initially, without considering the effects of any other variables, the four bendiness measures common to this study and the motivating study were calculated for rural crashes. All possible routes were weighted according to the intersection weighting method from Section 3.2.3. Results of the Fowler method were aggregated at a TLA level and compared with the results from the motivating study, as shown in the scatter plots of Figures 6.1 to 6.8. Figure 6.9 shows the cumulative frequency distributions of TLAs within certain percentage tolerances between the Haynes and Fowler methods for all measures. This gives a more quantitative understanding of the correlation between the two methods than the scatter plots.



**Figure 6.1 Bend Density Comparison by TLAs for Rural State Highways**



**Figure 6.2 Bend Density Comparison by TLAs for Rural non-State Highway Roads**

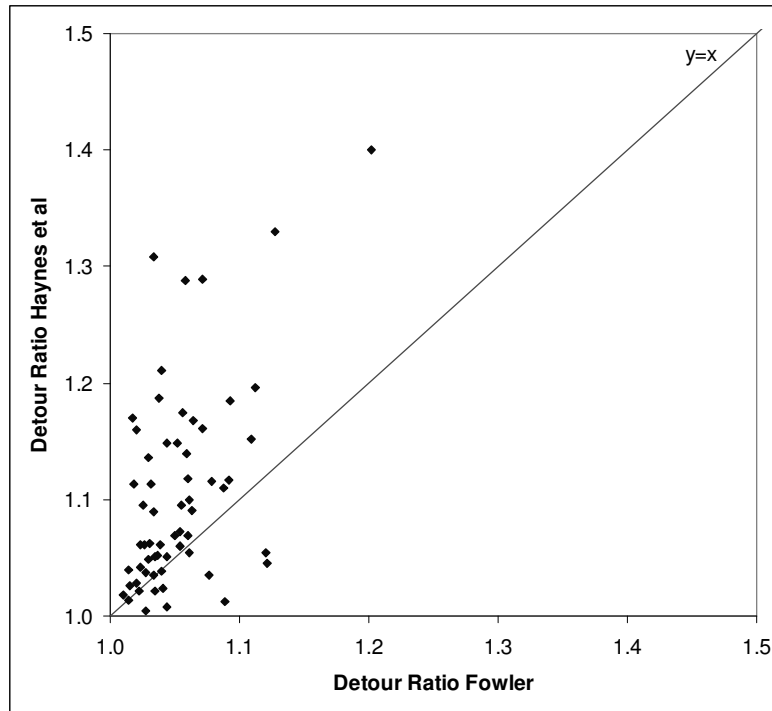


Figure 6.3 Detour Ratio Comparison by TLAs for Rural State Highways

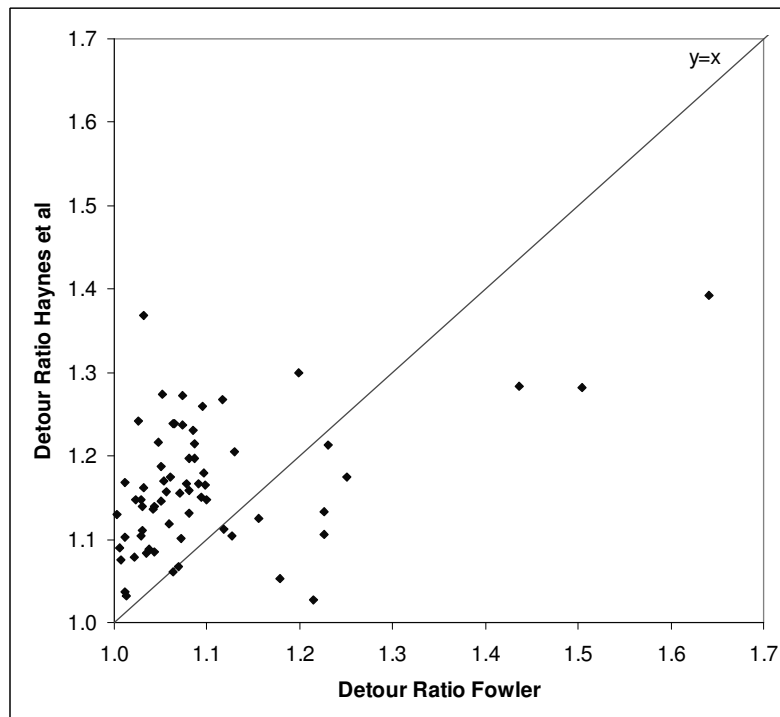


Figure 6.4 Detour Ratio Comparison by TLAs for Rural non-State Highway Roads

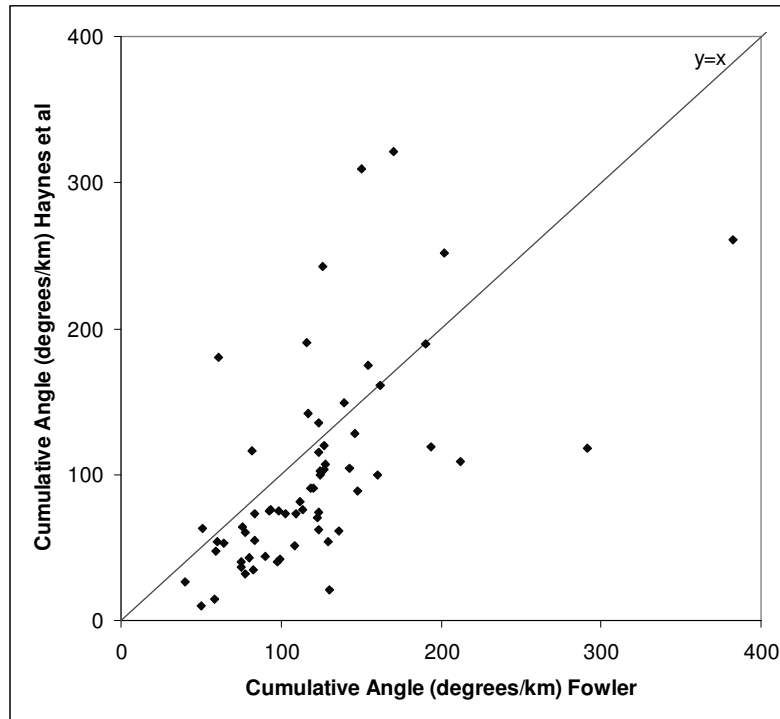


Figure 6.5 Cumulative Angle Comparison by TLAs for Rural State Highways

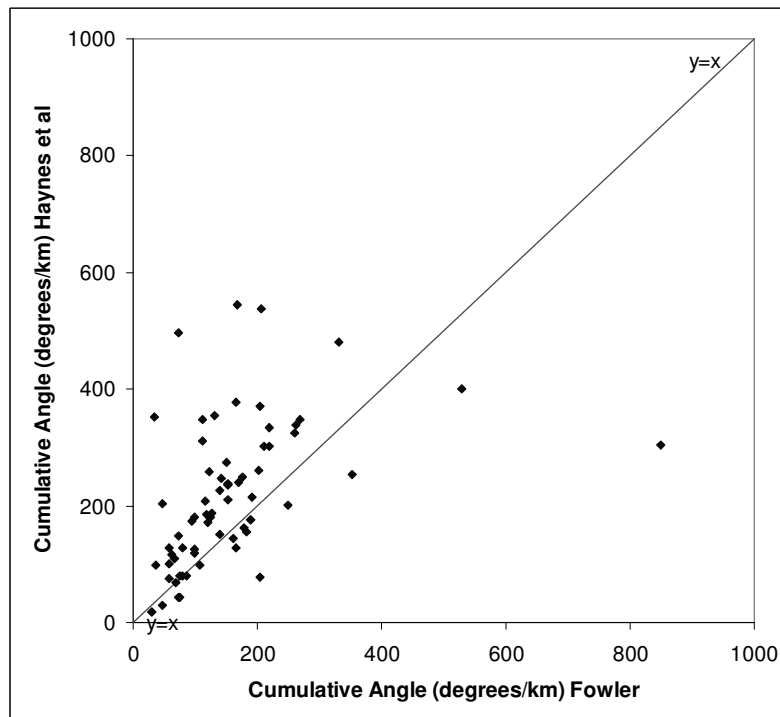


Figure 6.6 Cumulative Angle Comparison by TLAs for Rural non-State Highway Roads

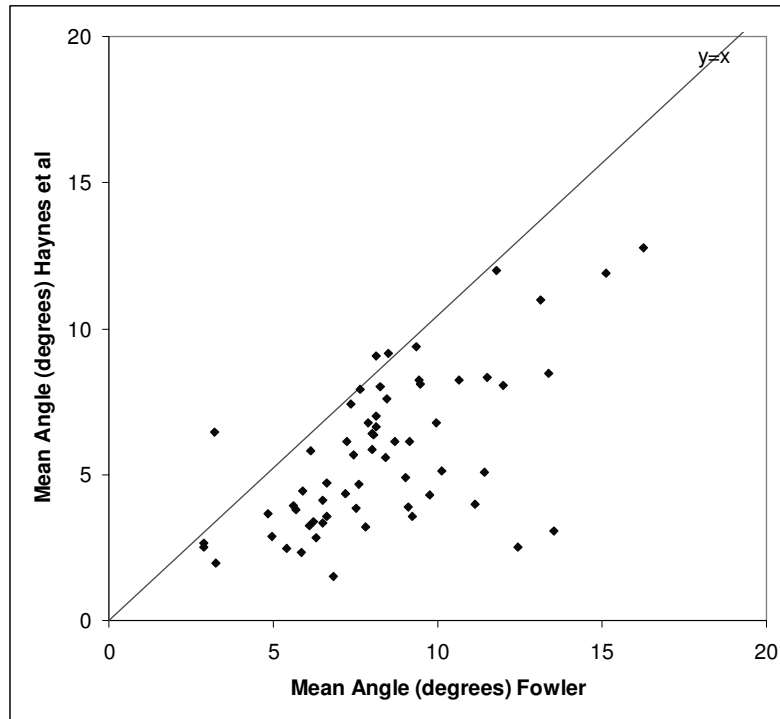


Figure 6.7 Mean Angle Comparison by TLAs for Rural State Highways

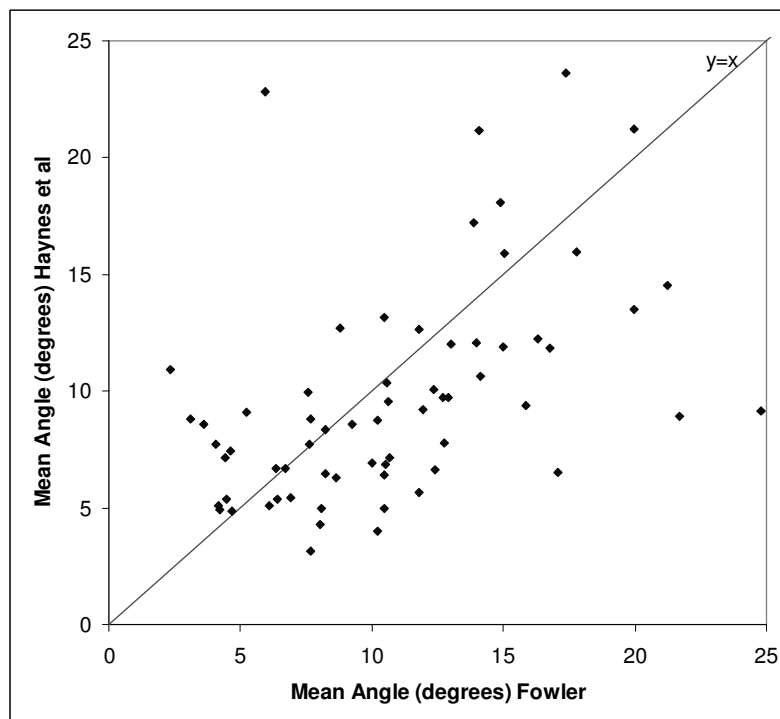
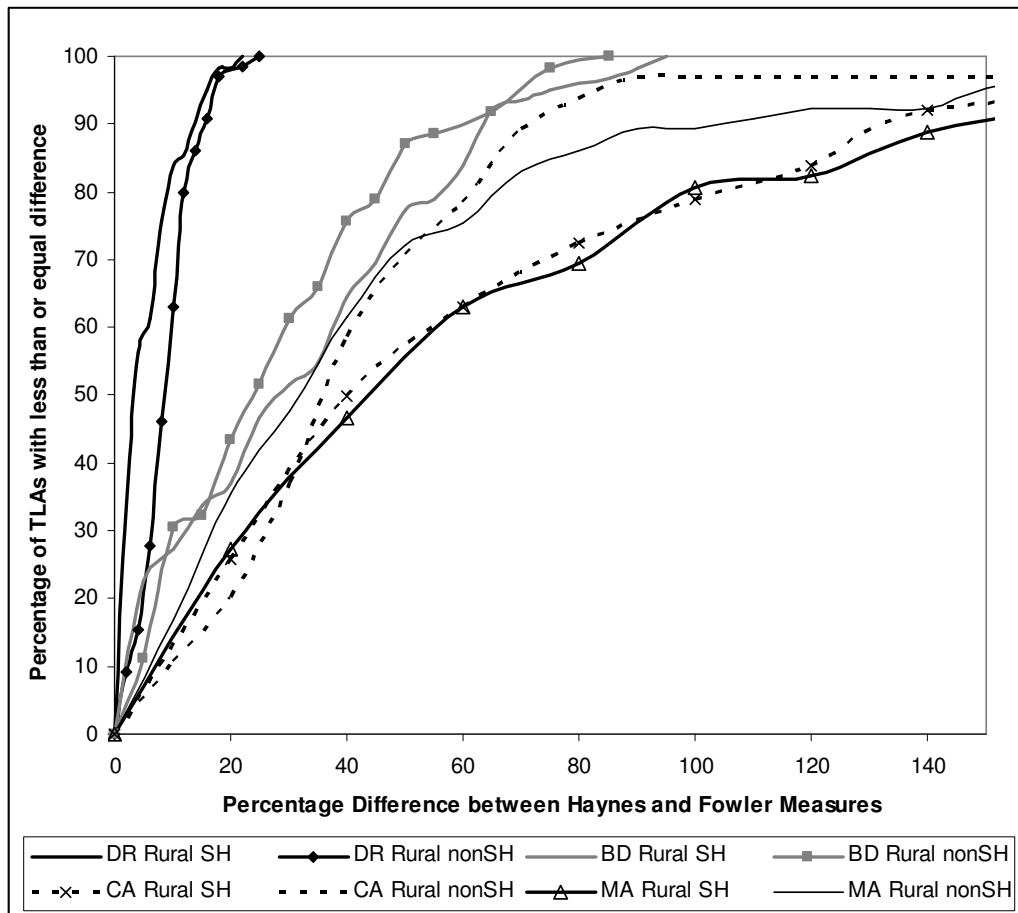


Figure 6.8 Mean Angle Comparison by TLAs for Rural non-State Highways



**Figure 6.9 Cumulative Frequency Distributions of Percentage Differences between Haynes and Fowler measures for Rural Roads**

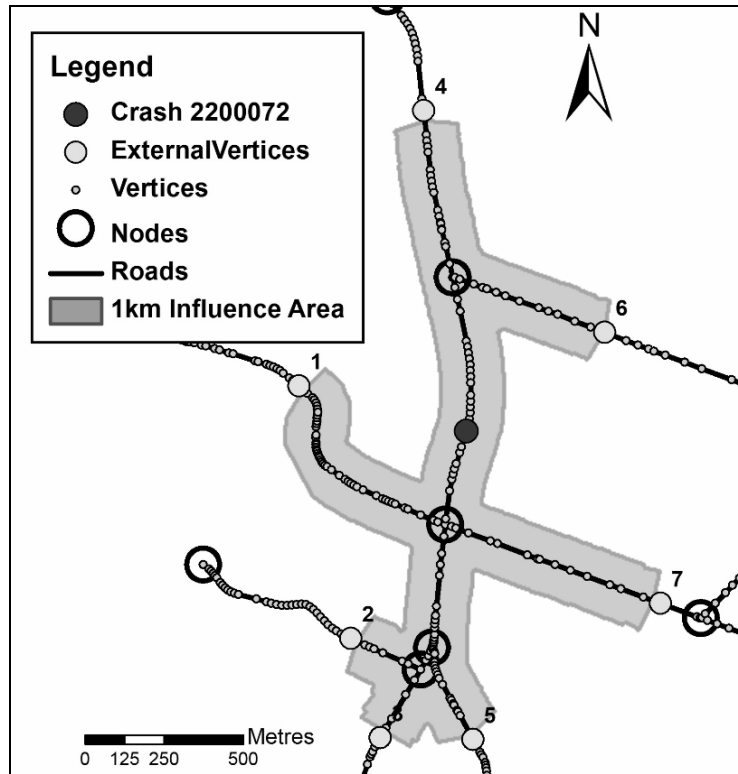
From the scatter plot comparisons (Figures 6.1 to 6.8) it appears that the bend density was the most consistent measure between the two studies, especially for rural state highways. Figure 6.9 also shows that the two methods produced well-correlated results with 47% and 52% of the Fowler method's rural state highway and non-state highway results respectively being within 25% of the Haynes results. Therefore it seems that averaging the bend density over a whole TLA area was a good approximation for the actual bend density of roads surrounding the crash location.

The cumulative angle comparisons also showed obvious correlations between the two studies, however the Haynes measures seemed to predict lower cumulative angle values than the Fowler measures for the state highways and greater values for the non-state highways. Figure 6.9 showed the rural non-state highway results to be especially poorly correlated. This could be in part due to the different resolutions of state

highway and non-state highway data. As the state highway parts of the network were generally the most detailed and the Haynes studies excluded any road links of length less than 200m (whereas the Fowler study did not) it would be expected that the Fowler cumulative angle values would be slightly higher.

The detour ratio values had similar ranges for both studies. The seemingly poor correlations shown in Figures 6.3 and 6.4 are contradicted by Figure 6.9 which shows the detour ratio to be the most consistent measure with 100% of the Fowler method's rural state highway and non-state highway results being within 25% of the Haynes results.

It should be noted that the Haynes methodology of calculating the detour ratio was adapted to suit the different sampling technique of the Fowler study. In order to calculate the detour ratio for the Fowler study points 1km of travel along the road from the crash site were considered to be nodes even though in most cases they weren't even vertices, just locations part-way along a link. Thus the route length was always 1km and the detour distance was calculated from the crash location (which was also often not actually a true node) to an exact point 1km away. For example, it can be seen in Figure 6.10 that neither the location of crash 2200072 or any of the external vertices at the edge of its influence area are actually nodes.



**Figure 6.10 Illustration of Nodes**

The mean angle plots were very dispersed with especially poor correlation between the two methods for the rural non-state highways, as shown in Figure 6.8. This poor correlation was confirmed by Figure 6.9 where the mean angle cumulative distributions were among the worst of all the measures. This was due to some of the TLAs having very high deviations between the Haynes and Fowler methods.

In general the state highway comparisons were more similar between the two methods than the non-state highway comparisons. This may be reflective of the different standards of data; state highway information is generally more accurate and provided at a higher resolution.

#### 6.1.2 Urban Road Crashes

Results of the Haynes and Fowler studies were compared in the same way for the urban roads of TLAs. Figures 6.11 to 6.14 give scatter plots where each point represents the Haynes and Fowler values of a measure for a particular TLA. Figure 6.15 shows the cumulative frequency distributions of TLAs within certain percentage tolerances between the Haynes and Fowler methods for all measures.



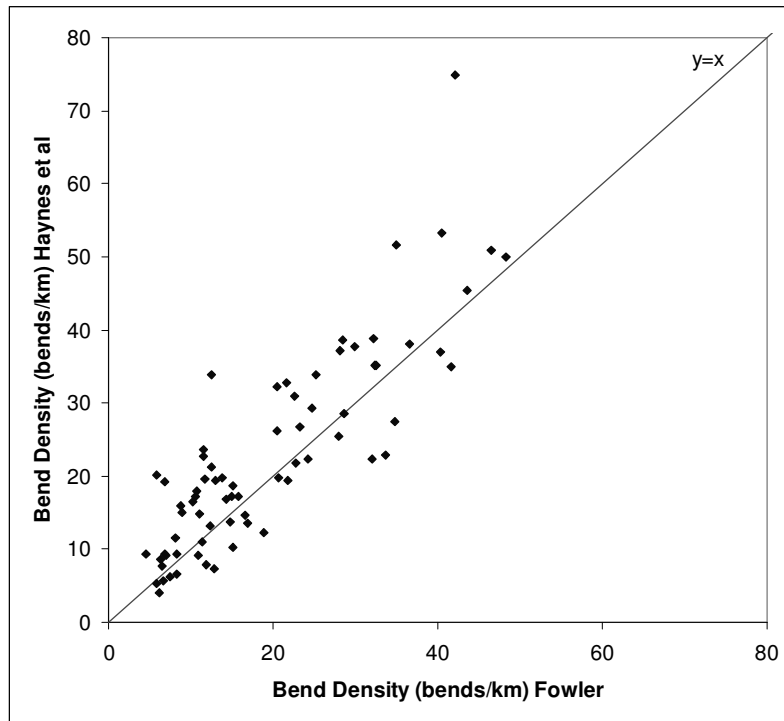


Figure 6.11 Bend Density Comparison by TLAs for Urban Roads

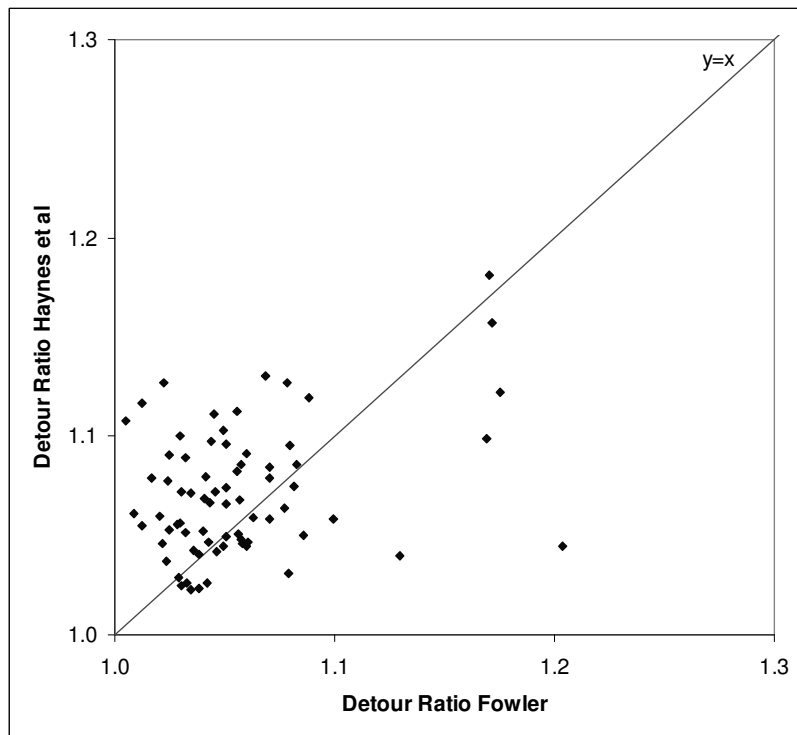
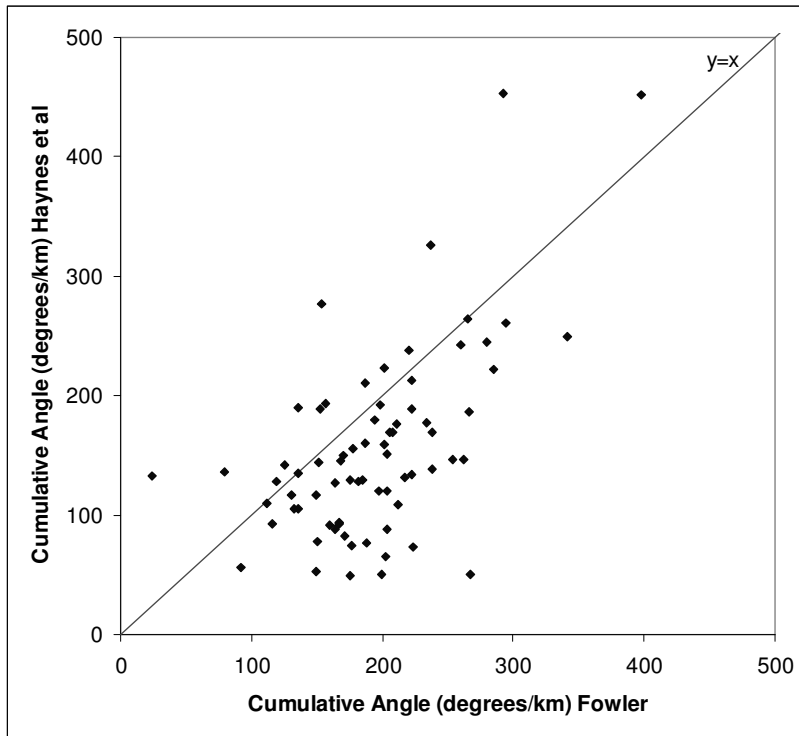
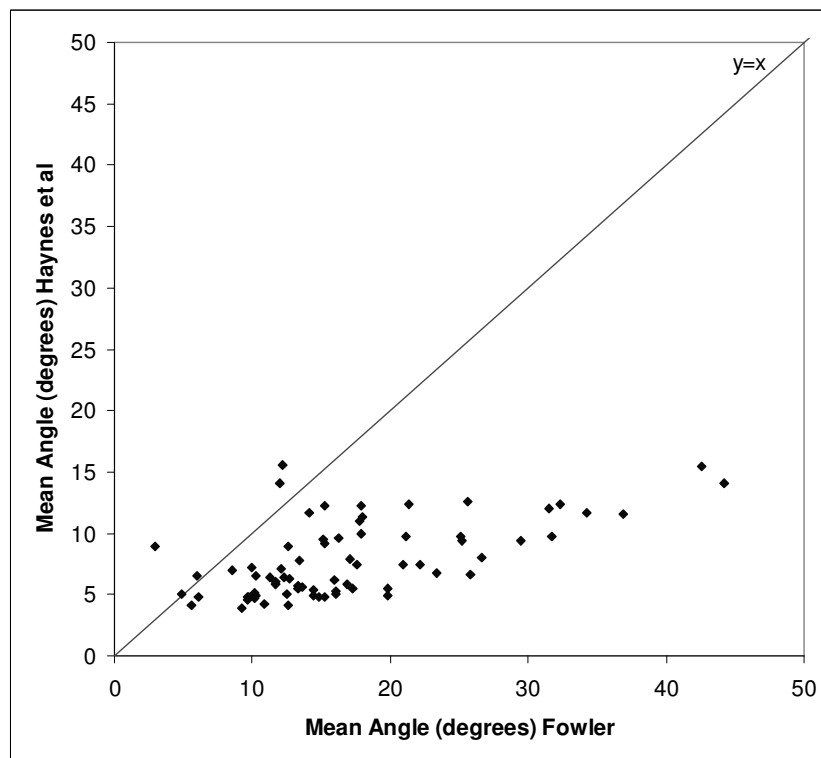


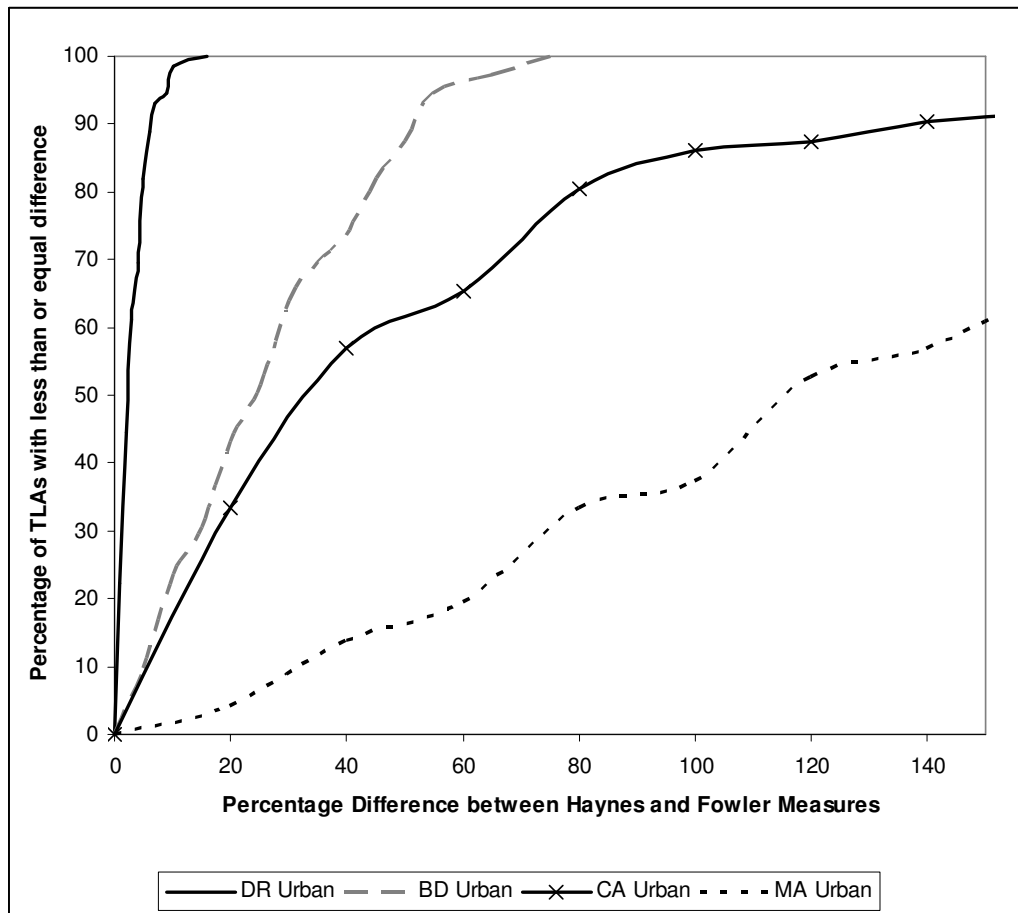
Figure 6.12 Detour Ratio Comparison by TLAs for Urban Roads



**Figure 6.13 Cumulative Angle Comparison by TLAs for Urban Roads**



**Figure 6.14 Mean Angle Comparison by TLAs for Urban Roads**



**Figure 6.15 Cumulative Frequency Distributions of Percentage Differences between Haynes and Fowler measures for Urban Roads**

In the urban road analysis, the bend density again seemed the most consistent measure between the two studies based on the scatter plots but the detour ratio had the best cumulative frequency distribution due to the small range of detour ratio values.

The high degree of scatter for the detour ratio (shown in Figure 6.12) would be expected in the urban situation where there is more variation of configuration within the network and a small sample is less likely to be representative of the whole area than for the rural situation. Also, as was postulated in the analysis of the rural results, some discrepancy was expected due to the slight differences between the two methods.

The mean angle comparison showed a definite trend between the two methods but, as this trend deviated from the  $y = x$  line the cumulative frequency distribution (Figure

6.15) was very poor. The Fowler method gave a much larger range of mean angle values than the Haynes method. This was probably because, for the urban situation, more turns at intersections are included in the Fowler method as it considered all possible routes originating from the location of interest rather than simply taking the shortest route between each node pair as the Haynes method did.

In general, there is enough deviation between the results of the two different methods to warrant trialling the Fowler method which uses a less aggregated approach and focuses on the influence areas of crashes rather than the characteristics of their entire TLA's network.

## **6.2 Non-Flow Weighted Comparisons between Crash and Non-Crash Sites**

Due to the method of aggregation used, the motivating study did not use separate comparison locations. Thus, while comparisons had been made between the Haynes results and the Fowler results at a TLA level of aggregation, it was considered important to also make an initial evaluation of the properties of the comparison locations with respect to the crash locations before further development of the Fowler method. To do this, the distributions of the initial bendiness measure calculations (without consideration of traffic flow or other influencing variables) were compared between the crash and non-crash sites for the rural state highway, rural non-state highway and urban cases.

### **6.2.1 Rural Roads**

Distributions of the bendiness measures for actual crash locations and non-crash comparison locations situated on rural state highways are shown in Figures 6.16 to 6.20:

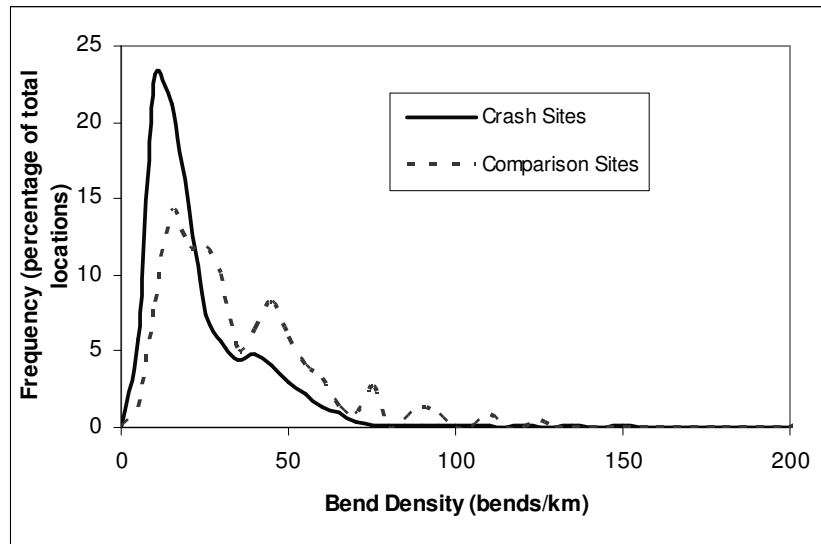


Figure 6.16 Bend Density for Rural State Highway Crash and Comparison Sites

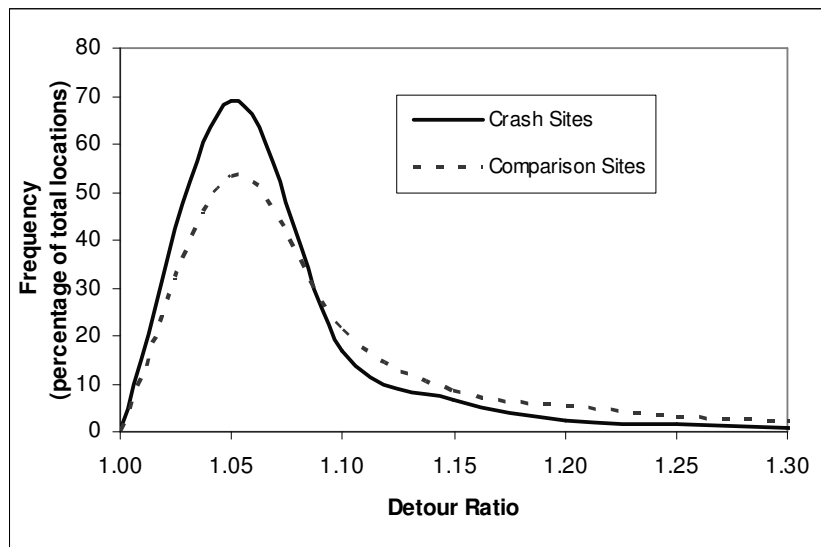


Figure 6.17 Detour Ratio for Rural State Highway Crash and Comparison Sites

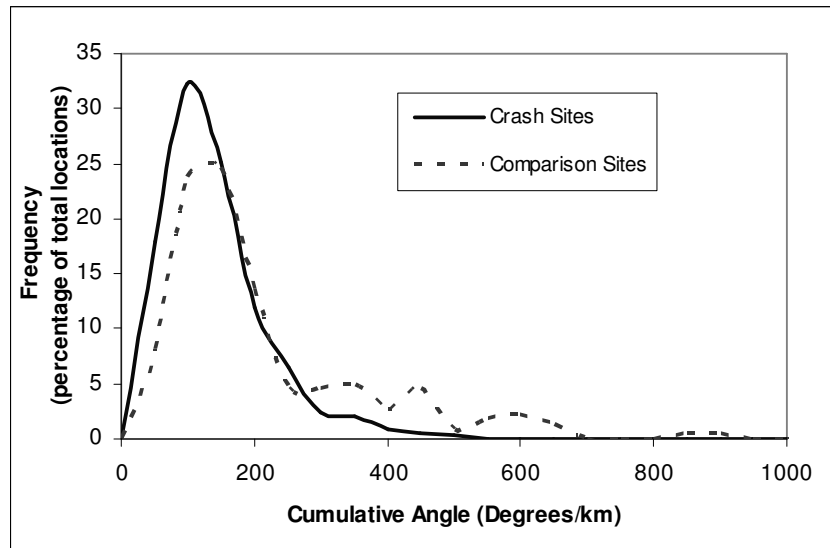


Figure 6.18 Cumulative Angle for Rural State Highway Crash and Comparison Sites

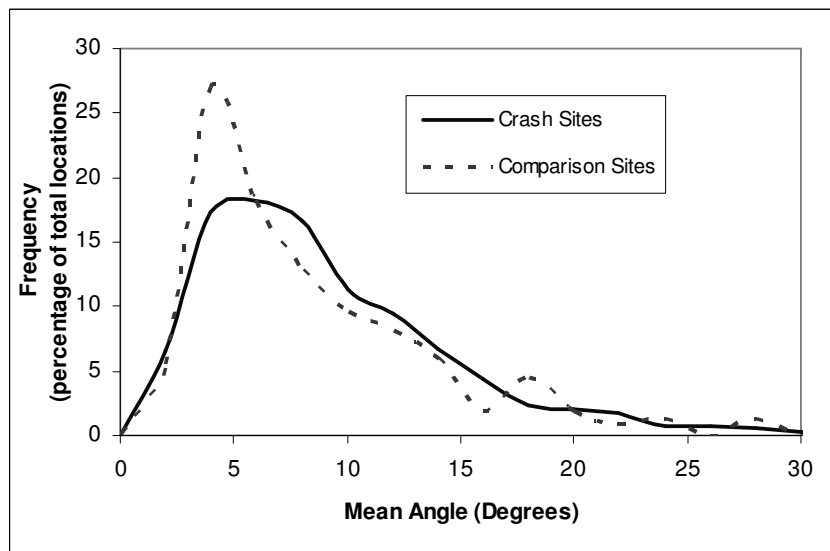
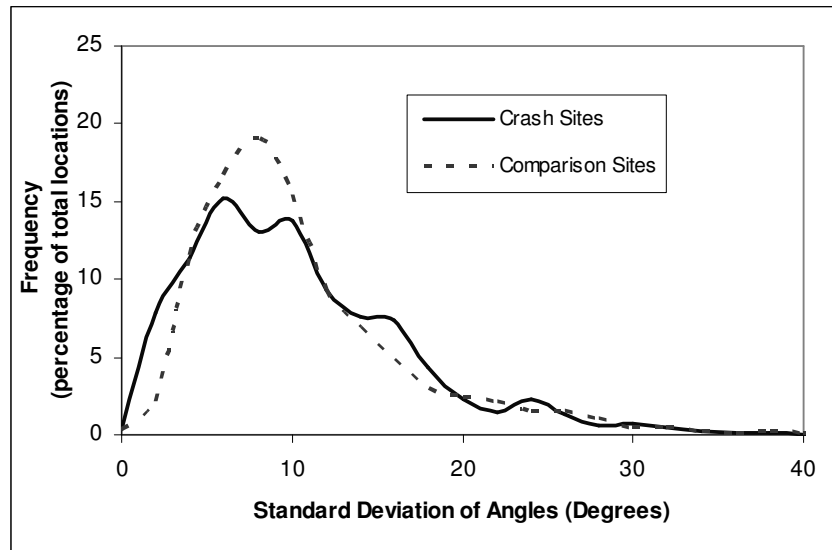


Figure 6.19 Mean Angle for Rural State Highway Crash and Comparison Sites

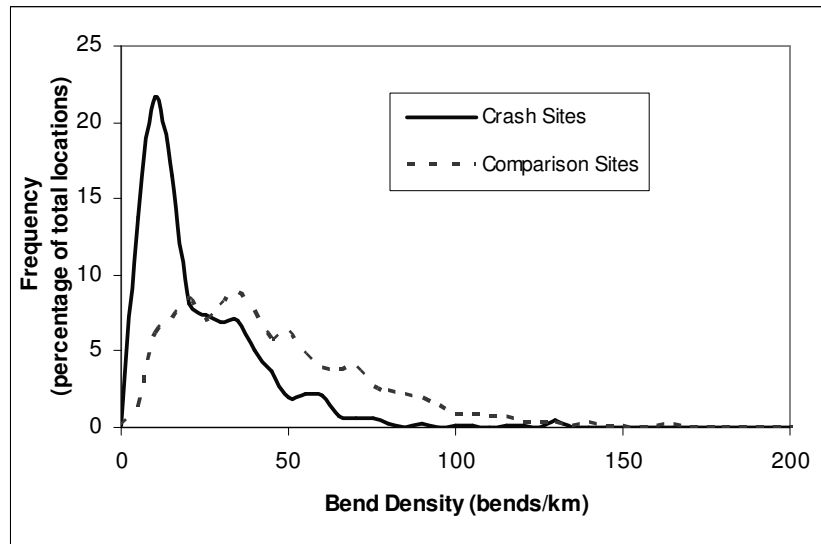


**Figure 6.20 Standard Deviation of Angles for Rural State Highway Crash and Comparison Sites**

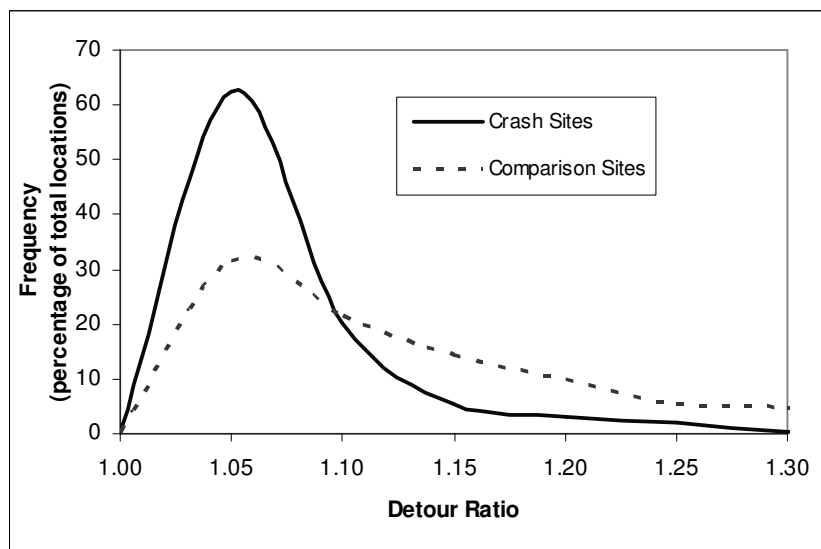
For the rural state highway locations the five measures had very similar distributions with very similar mean values between crash and comparison sites, which probably indicates the consistency of design of New Zealand's state highways. The distributions of the three measures that incorporate distance (bend density, detour ratio and cumulative angle, shown in Figures 6.16 to 6.18) had much smaller spread for the crash sites than the comparison sites. This suggests that crashes are less likely to occur in regions of higher bendiness. The comparison site distribution has a lower spread than the crash site distribution for the mean angle (Figure 6.19) but still has a higher average value, which supports the theory that bendiness is protective.

The standard deviation of angles distributions (Figure 6.20) showed lower mean values and slightly lower spreads for the comparison sites. This suggests that it is important to have a high level of design consistency.

The corresponding comparisons for the bendiness measures of crashes and comparison locations on rural non-state highway roads are shown in Figures 6.21 to 6.25:



**Figure 6.21 Bend Density for Rural Non-State Highway Crash and Comparison Sites**



**Figure 6.22 Detour Ratio for Rural Non-State Highway Crash and Comparison Sites**



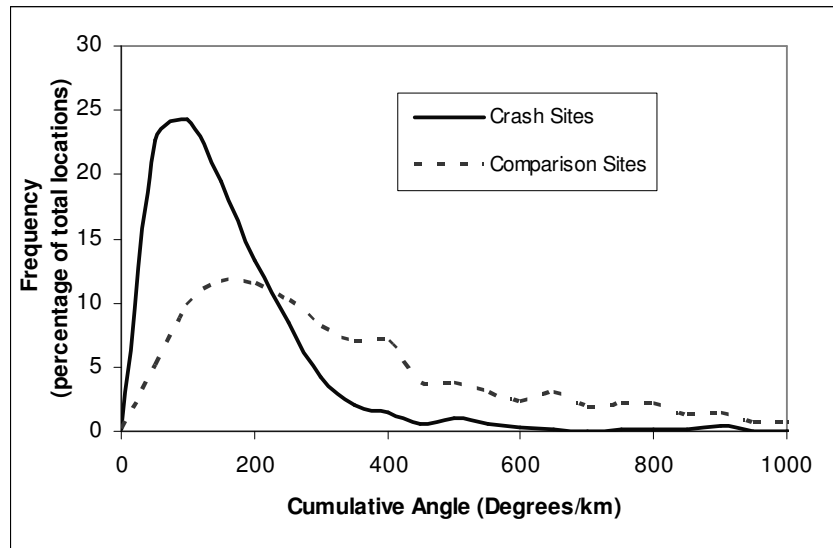


Figure 6.23 Cumulative Angle for Rural Non-State Highway Crash and Comparison Sites

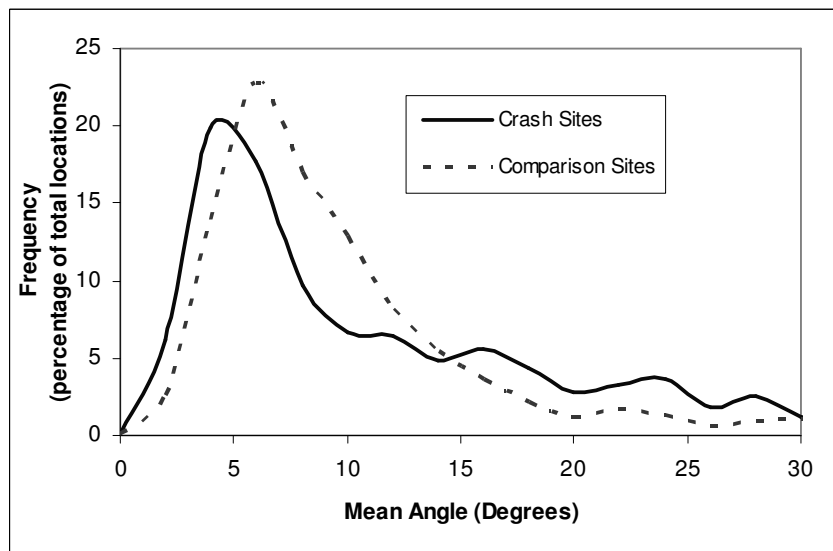
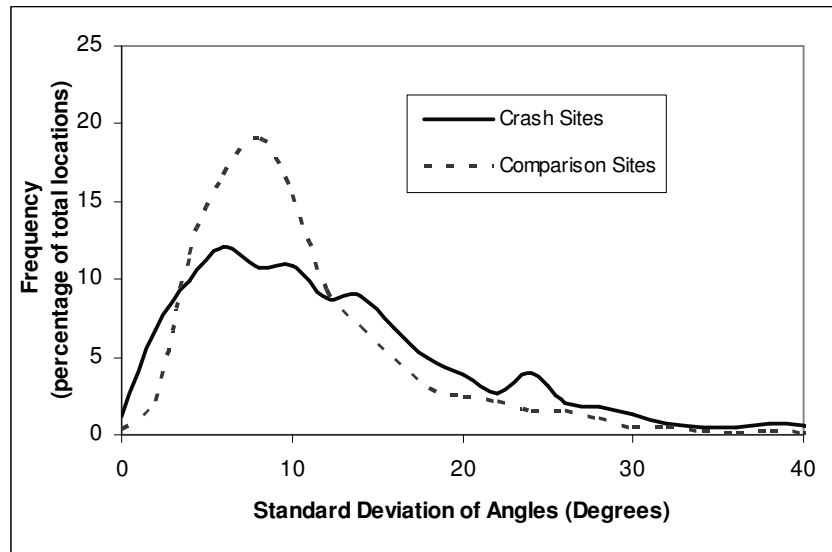


Figure 6.24 Mean Angle for Rural Non-State Highway Crash and Comparison Sites



**Figure 6.25 Standard Deviation of Angles for Rural Non-State Highway Crash and Comparison Sites**

The non-state highway rural crash and comparison site distributions showed similar trends to the rural state highway sites but with much greater distinction. The crash sites generally had lower mean values with lower spreads, which again indicates that bendiness is protective. In terms of the standard deviation of angles measure, intended to indicate the consistency of routes (Figure 6.25), the crash distribution had a higher mean and higher spread, which again indicates that design consistency is protective.

### 6.2.2 Urban Roads

As for the rural sites, the bendiness measure distributions of crash and comparison locations situated on urban roads were also computed. These are shown in Figures 6.26 to 6.30:

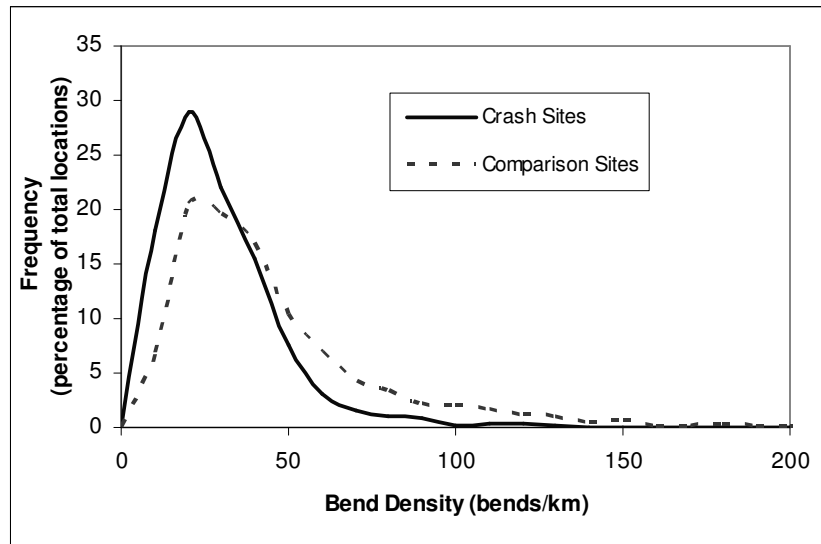


Figure 6.26 Bend Density for Urban Crash and Comparison Sites

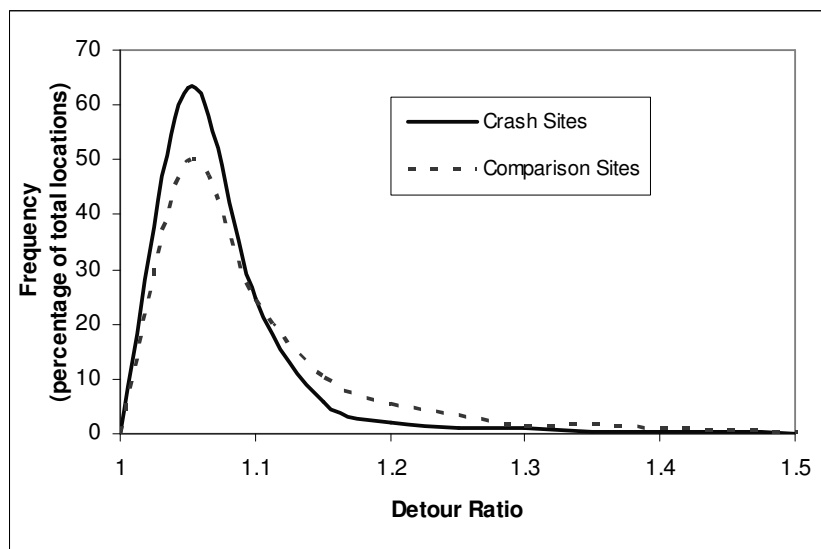


Figure 6.27 Detour Ratio for Urban Crash and Comparison Sites

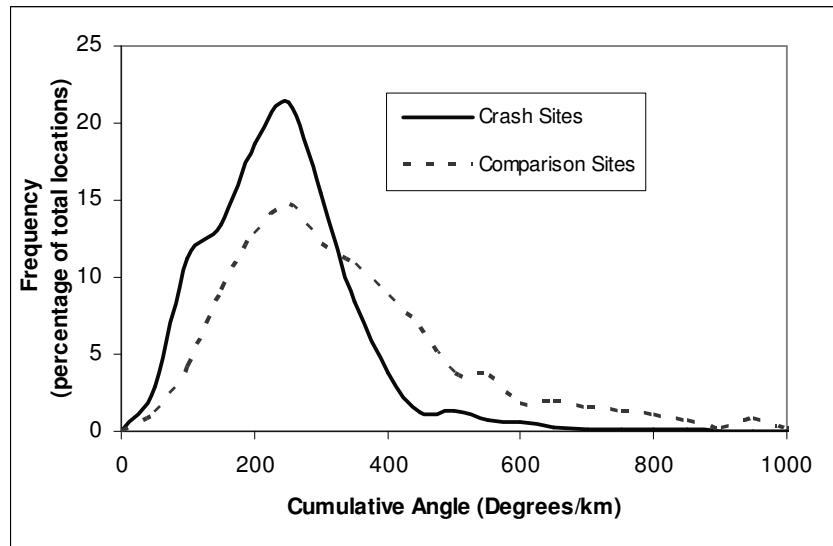


Figure 6.28 Cumulative Angle for Urban Crash and Comparison Sites

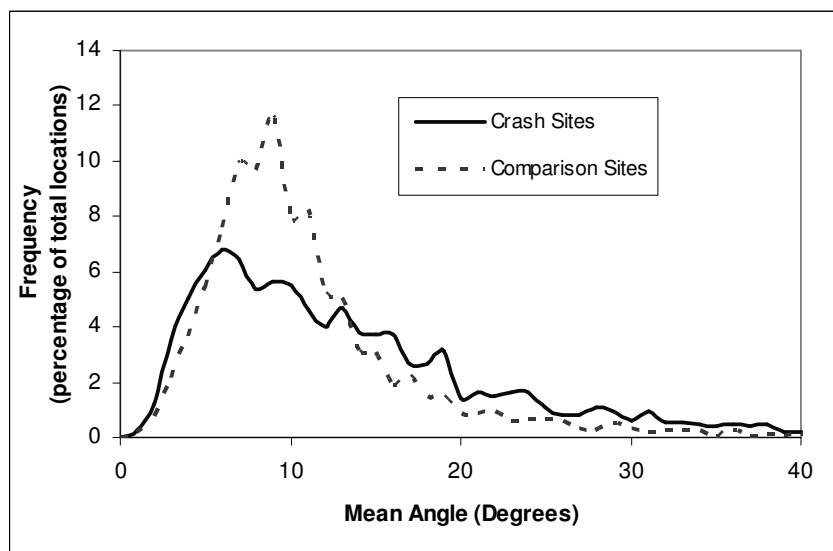
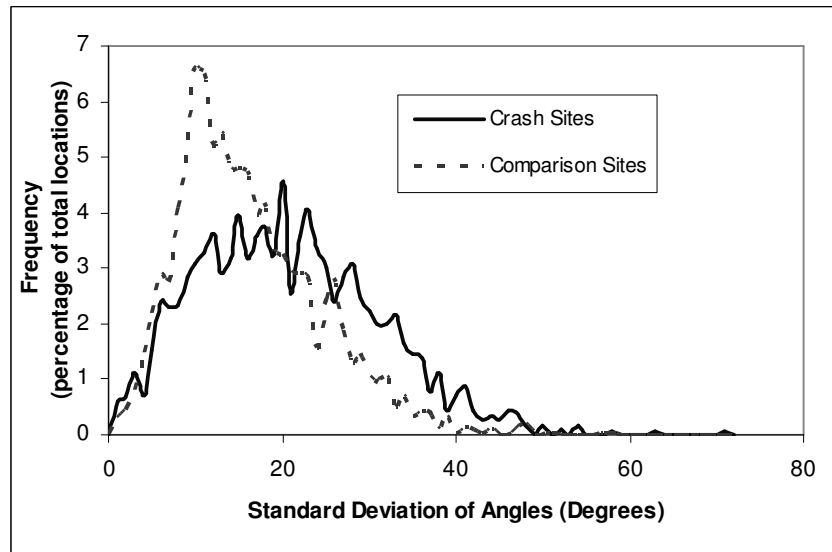


Figure 6.29 Mean Angle for Urban Crash and Comparison Sites



**Figure 6.30 Standard Deviation of Angles for Urban Crash and Comparison Sites**

The urban distributions showed similar trends to the rural ones, especially for the bend density, detour ratio and cumulative angle measures (Figures 6.26 to 6.28), although the distinctions between urban crash and comparison sites did not seem as pronounced as for the rural non-state highway sites. Mean values were much higher for the urban distributions than the rural ones, which would be expected given that urban locations have much greater junction densities.

The standard deviation between angles distributions (Figure 6.30) showed a more pronounced difference for the urban case; the crash sites seemed to have higher deviations between successive elements. This might imply that design consistency is especially important in the urban environment. The difference could also be due to a difference in junction densities between crash and comparison sites, which would imply that the comparison sites were not appropriately chosen.

It should be noted that, while these comparisons may give useful indications of the differences between crash and non-crash sites, the results should not be seen as conclusive. Many factors other than bendiness that affect crash occurrence exist but were not accounted for in these simple comparisons, the most important being traffic flow through the sites. Some attempt was made to match the crash sites with similar comparison sites by separating the rural state highway and other rural road sites. It

would be more appropriate, however, if comprehensive flow data were available, to use flow ranges to identify sites of similar exposure for comparison.

Thus the next section of this thesis, which details the estimation of models that considered many possible influencing variables, including those intended to represent traffic flow, is an example of a much more rigorous investigation.

### 6.3 Bendiness Measures applied with Flow Weighting in Upper Hutt Region

As described previously, a comprehensive flow data set that covered all road types within a region was only available for the Upper Hutt City TLA. Thus the bendiness measures relating to the majority of sites throughout New Zealand could only be analysed using the non-flow weighted methods. 15 of the crashes from the whole crash sample occurred in the Upper Hutt City region, as shown in Figure 6.31:

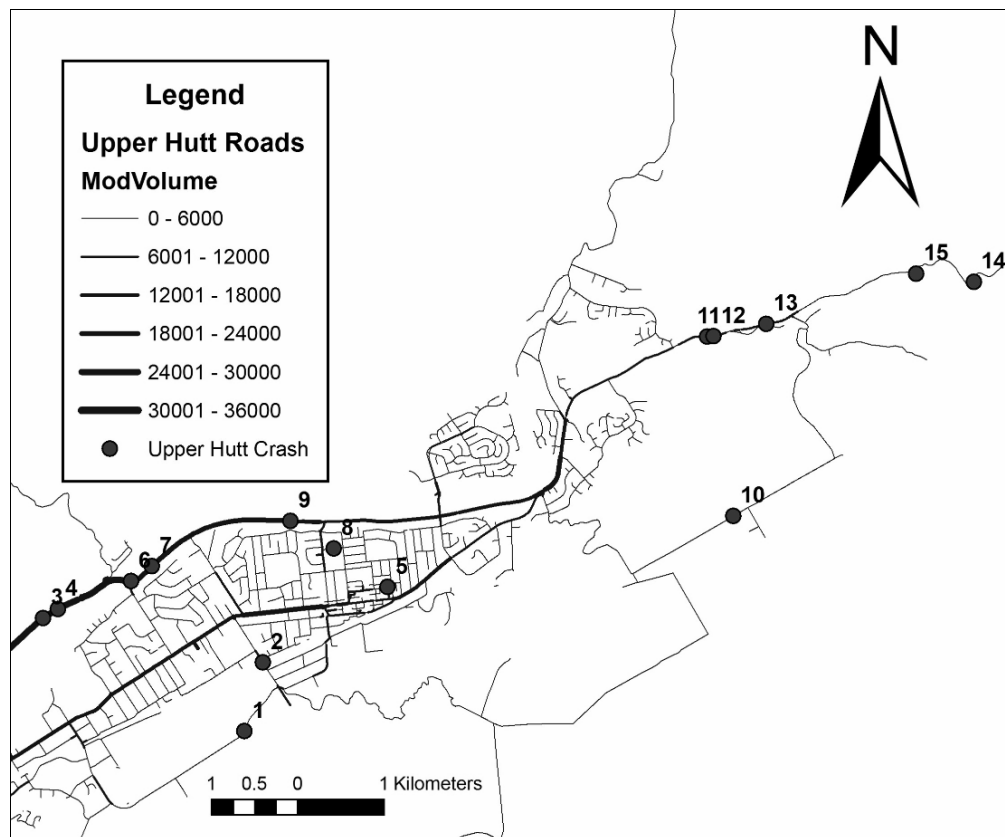


Figure 6.31 Upper Hutt Crashes

The nature of the Upper Hutt region's road hierarchy can also be seen from Figure 6.31. Major arterial roads with high traffic flows are connected to local roads with small flows by collector roads. As expected, due to the greater exposure to crash risk on arterial roads, the majority of crashes seem to occur on the arterial roads.

The Upper Hutt City information was used to test the appropriateness of using non-flow weighted methods where flow data were not available. Comparisons for each of the bendiness measures applied to the 15 crashes in the Upper Hutt region were made between the non-flow weighted (Section 3.2.3) and flow weighted methods (Section 3.2.4). Comparison plots and their associated Pearson correlation coefficients are shown in Figures 6.32 to 6.36:

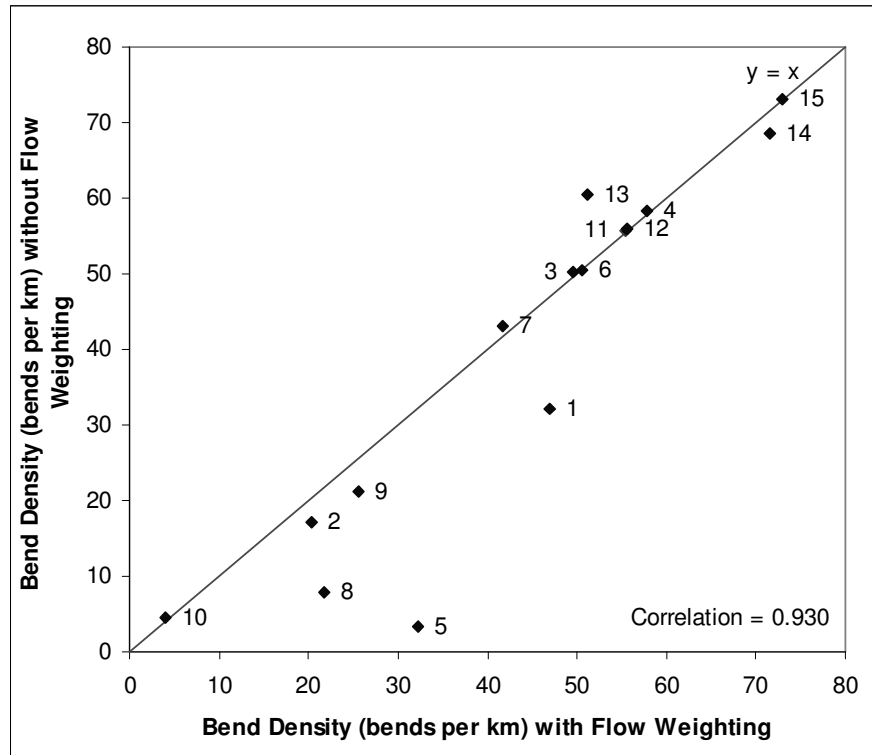


Figure 6.32 Bend Density Results for Upper Hutt Crashes with and without Flow Weighting

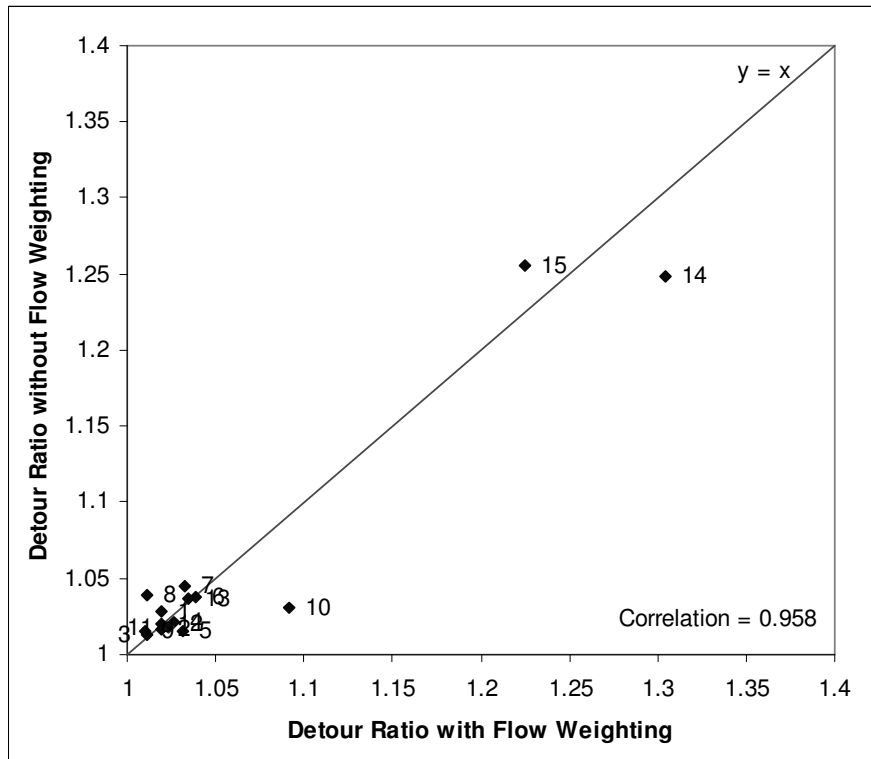


Figure 6.33 Detour Ratio Results for Upper Hutt Crashes with and without Flow Weighting

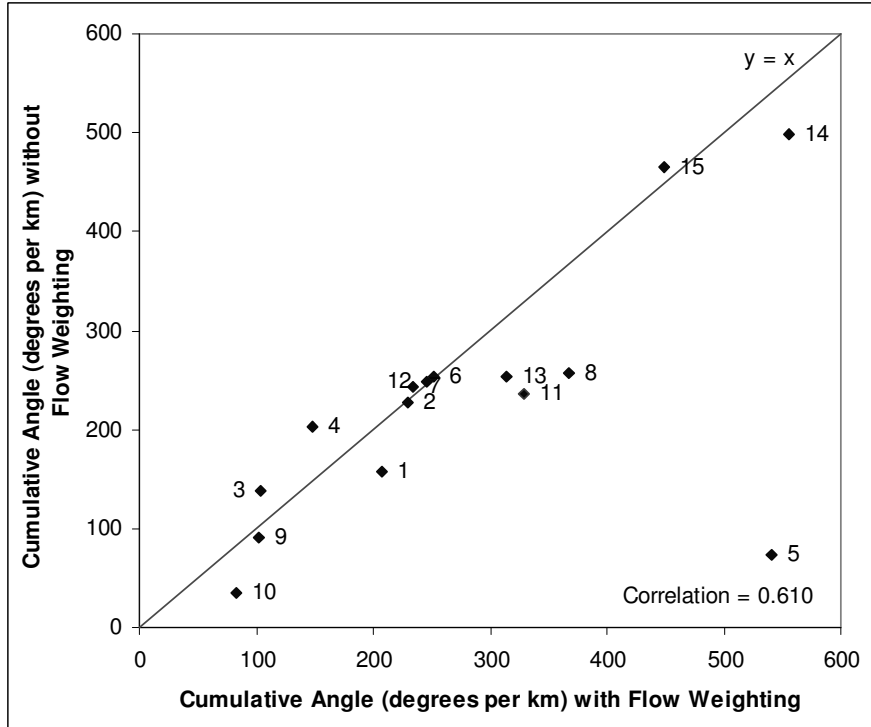


Figure 6.34 Cumulative Angle Results for Upper Hutt Crashes with and without Flow Weighting



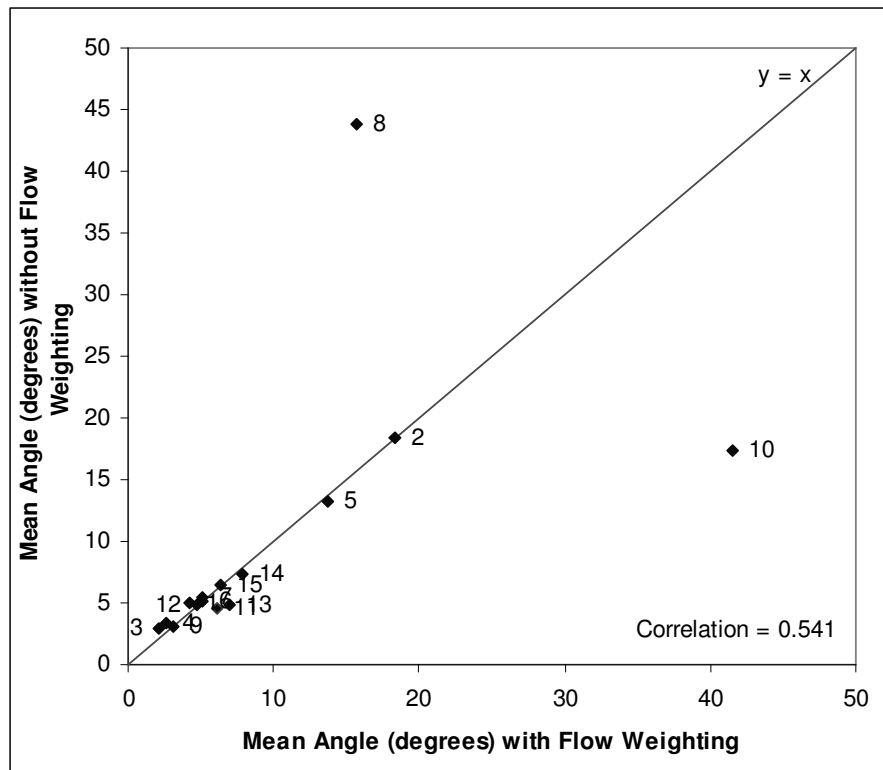


Figure 6.35 Mean Angle Results for Upper Hutt Crashes with and without Flow Weighting

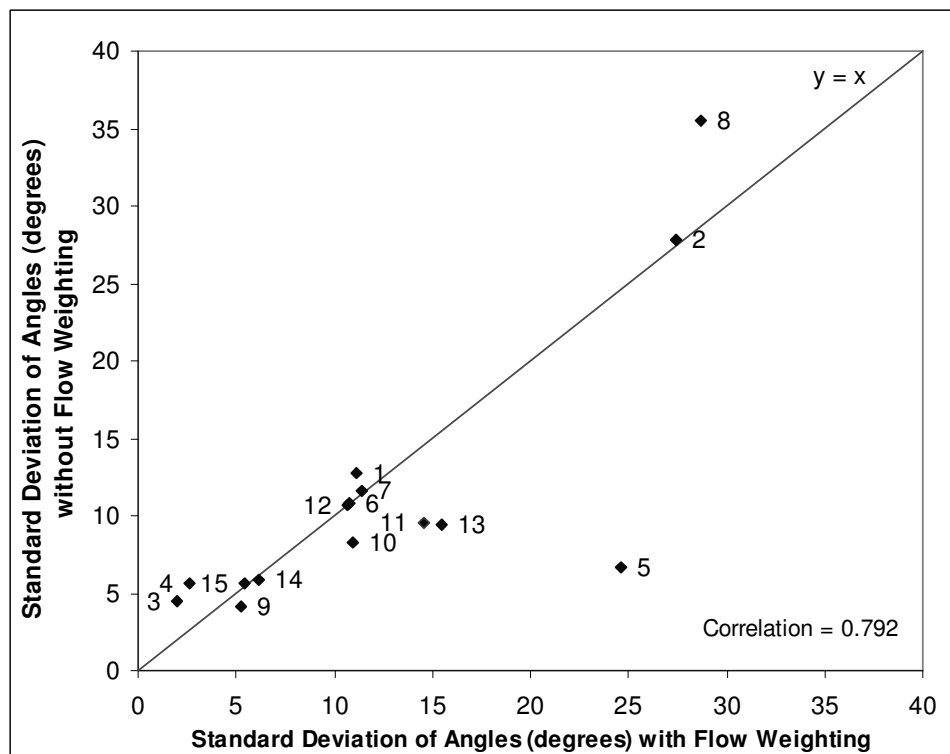


Figure 6.36 Standard Deviation of Angles Results for Upper Hutt Crashes with and without Flow Weighting

It can be seen that in most cases the two methods gave very similar results. The bend density and detour ratio measures had much better correlations than the measures that involved angles. The standard deviation of angles term also had a reasonably high correlation between the flow weighted and non-flow weighted cases. This was good as this measure was the only one that represented design consistency.

Outliers were not always the same for each measure but it was noted crashes 5, 8 and 10 had consistently poor comparisons between the two measures. These crashes were situated in areas of high variability of flow between surrounding links. This may not be obvious for crash number 10 but closer inspection shows that one of the links was predicted by the model to have negligible flow.

Thus, as expected from comparisons of the two weighting equations, the non-flow weighting method gives a reasonable approximation for bendiness measures in areas with small variability of flow. Applying the non-flow weighting method to areas with distinct road hierarchies and large flow variations is not recommended. However, it can be assumed that for one kilometre stretches of road such situations will not be too frequent. Thus, the findings of this investigation warrant using the non-flow weighted data for further analysis. It was determined that, given the small sample size of the Upper Hutt City crash data, there would not be much point in creating a separate model to incorporate the flow data of the region so the Upper Hutt City data was analysed with the rest of the country's data according to the non-flow weighted method.

## 7.0 Case Study Results

This section presents the final results for the case study. It consists of three main parts. Firstly, binary logistic regression models relating crash occurrence to bendiness, consistency and other influencing variables were created for three road types. Next, a method of comparing the consistency of bendiness measures over immediate and influence areas was tested.

### 7.1 Model Estimation

Models were developed for the rural state highway, rural non-state highway and urban roads cases. As detailed in Section 3.3.2, binary logistic regression was used. Several bendiness measures were used in each model to give a more complete view of bendiness. Other influencing variables were also tested.

Variables were chosen to be part of a model based on the following criteria:

- Their individual correlations with crash occurrence (desired a high absolute Pearson product moment correlation coefficient with low associated p-value). Correlation tables for the three models are shown in Appendix 11.2;
- Their individual correlations with other predictor variables (desired low correlations with low associated p-values between predictor variables);
- Their p-value when included in the model; and
- The effect they had on the general model evaluation statistics (number of concordant pairs, chi-squared statistics etc) when included in the model.

#### 7.1.1 Rural State Highway Crash Model

The model chosen to predict crashes on rural state highways is summarised in Equation 24 and Table 7.1:

$$\log it[Y] = -1.49 + 2.60DR - 6.72CA + 0.0757SD + 0.000328DT - 0.000984ME$$

#### **Equation 24 Rural State Highway Model**

where:  $Y$  = the probability of a crash occurring ( $0 \leq Y \leq 1$ )

$DR$  = Detour Ratio

$CA$  = Cumulative Angle (degrees/m)

$SD$  = Standard Deviation of Angles (degrees)

*DT* = Average Annual Daily Traffic (estimated by Transit NZ)

*ME* = Mean Elevation above sea level (metres)

**Table 7.1 Rural State Highway Model**

<b>Predictor</b>	<b>Coefficient</b>	<b>Standard Error of Coefficient</b>	<b>P-value</b>
<b>Constant</b>	-1.49	1.28	0.243
<b>Detour Ratio</b>	2.60	1.26	0.039
<b>Cumulative Angle (degrees/m)</b>	-6.72	0.95	<0.001
<b>Standard Deviation of Angles (degrees)</b>	0.0757	0.0163	<0.001
<b>Estimated Average Annual Daily Traffic (veh)</b>	0.000328	0.000040	<0.001
<b>Mean elevation above sea level (m)</b>	-0.000984	0.000377	0.009

The evaluation statistics of the model were as follows:

Test that all slopes are zero:  $G = 269$ , P-Value <0.001

Deviance goodness of fit test:  $\chi^2 = 1044$ , P-Value > 0.999

Percentage of pairs that were concordant = 81.3%

Percentage of pairs that were discordant = 18.3%

Percentage of pairs that were tied = 0.4%

In terms of choosing which of the bendiness variables to use, the bend density and cumulative angle measures were highly correlated; cumulative angle proved to be the most effective when included in the model so bend density was excluded. Similarly, the mean angle and standard deviation of angles measures were highly correlated; standard deviation proved to be the most effective when included in the model and added an indication of the effect of design consistency, so mean angle was excluded. Detour ratio also had a high correlation with cumulative angle, but had a small enough p-value and large enough effect on the model's performance to retain it in the model.

The estimated average annual daily traffic count at the locations of interest had a definite impact on the model's performance. It was interesting to note that this measure had a very low correlation with the measure expected to be a proxy for traffic flow (the sum of people residing and employed in the meshblock divided by its area) but a much higher correlation with the meshblock area itself. Although this did not necessarily mean that meshblock area would be the most appropriate proxy measure for rural non-state highway and urban roads (as they have different characteristics) it provided a useful starting point.

The mean elevation term had a significant correlation with crash occurrence and a negative correlation with traffic flow. This would be expected given the nature of the New Zealand road network where roads located in mountainous regions generally have lower flows. Elevation was positively correlated to all the bendiness measures except standard deviation of angles showing that mountainous roads are generally more bendy than roads at lower elevations, as would be expected considering differences in terrain.

The standard deviation of elevations term did not enhance the model's performance and therefore was not included. This suggests that consistency of a road's vertical alignment is not as important as consistency of its horizontal alignment on rural state highways, or that the vertical alignment of state highways is already consistent between sections.

The rainfall measures also did not appear to be significant when included in the model. This suggests that weather conditions do not affect crash occurrence on rural state highways (or that measures taken by traffic engineers to improve visibility and skid resistance etc. in wet areas are effective).

The measure used to distinguish whether or not a crash was within 30m of an intersection was also not significant. The vast majority (94%) of crashes considered were not close to intersections, thus it is likely that characteristics of non-intersection crashes dominated the model and there was not enough intersection crashes to warrant an alternative model.

Both junction density measures had good individual correlations with crash probability but did not improve the model or have significant p-values when included so were excluded to reduce the amount of data required.

**Table 7.2 Summary of Influencing Variables for Rural State Highway Model**

<b>Variable</b>	<b>Range</b>	<b>Mean</b>	<b>Coefficient</b>	<b>Coeff*Range</b>	<b>Coeff*Mean</b>
<b>Detour Ratio</b>	1; 1.91	1.06	2.60	2.60; 4.96	2.75
<b>Cumulative Angle (degrees/m)</b>	0; 1.35	0.13	-6.72	0.00; -9.10	-0.85
<b>Standard Deviation of Angles (deg)</b>	0; 44.4	9.27	0.0757	0.00; 3.36	0.70
<b>Estimated Average Annual Daily Traffic (veh)</b>	129; 24653	5409	0.000328	0.04; 8.09	1.78
<b>Elevation above sea level (m)</b>	0.7; 1056	152	-0.000984	0.00; -1.04	-0.15

The model showed that the probability of crash was directly related to the detour ratio. The greater the distance travelled compared to the direct distance between nodes, the greater the crash risk. The probability of crash was negatively related to cumulative angle. The greater the angle travelled, the lower the crash risk. The probability of crash was directly related to standard deviation of angles. The more variation between angles travelled (lower consistency), the higher the crash risk.

The results for detour ratio and cumulative angle may seem contradictory. Both had a negative relationship with probability of crash when individually correlated, indicating that bendiness is a protective quality. However, in this model with other factors accounted for, the detour ratio variable was the more influential, based on the

factors' model coefficients multiplied by mean values as shown in Table 7.2. This suggests that bendiness is a hazardous quality and that crash occurrence increases with increased bendiness on rural state highway roads. To explore this further, an analysis of typical road sections is presented later on.

The standard deviation of angles term had a positive correlation when compared to crashes and a positive coefficient in this model. This suggests that as consistency between elements increases the probability of a crash decreases. This effect, however, is much less important than that of the detour ratio, as shown in Table 7.2.

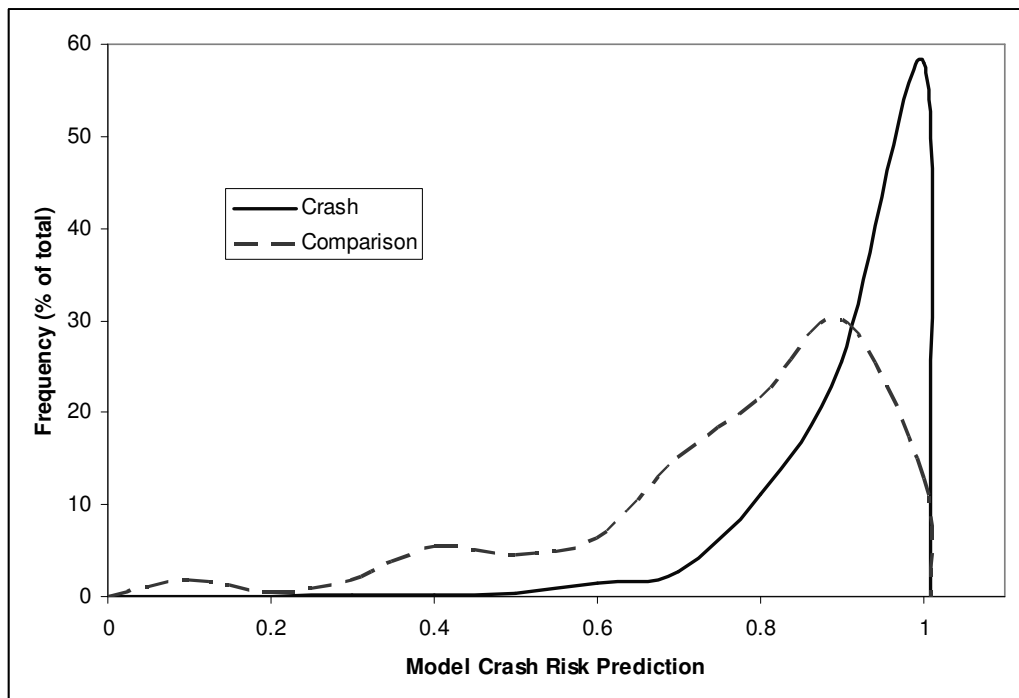
In terms of other, non-bendiness variables, the estimated traffic flow at each location of interest was shown to be very influential on the model, second only to the detour ratio value. The higher the flow along the road, the greater the crash risk, as would be expected given that higher flows mean greater exposure to the possibility of crashing.

In terms of the statistical evaluations, the p-values for all influencing variables included in the model were well below the level of significance, which suggests that these coefficients are not zero and therefore the variables were appropriately chosen. This is reinforced by the test that all slopes are zero, which had a reasonably low G-value and very low p-value indicating that at least one of the coefficients is not equal to zero. The p-value for the constant term was above the level of significance and the coefficient was very low, which indicates that there is a chance that the constant is equal to zero and therefore no important influencing variables have been excluded from the model.

The deviance chi-squared goodness-of-fit statistic had the highest possible p-value which would suggest that there was no evidence to reject the null hypothesis of the model being a good fit. Similarly, the number of concordant pairs was 81.3% which is very high and shows that the model is very good at distinguishing between locations of differing crash probabilities.

Figure 7.1 shows the distributions of crash risk for crash and comparison sites predicted by the model. Only one crash location was given a crash risk of less than 0.2; this occurred on a particularly bendy stretch of state highway. However, 43% of

comparison locations were given crash risks greater than 0.8, these sites generally had lower than average detour ratio and cumulative angle values and higher than average standard deviation of angles and AADT values. This suggests that the model is prone to over-predicting crash risks for regions of low bendiness and poor design consistency and such areas should be treated with caution. As a general rule, the model should not be trusted for analysis of sections with  $DR < 1.04$ ,  $CA < 0.10$  deg/m and  $SD > 8.6$  deg.



**Figure 7.1 Rural State Highway Model Crash Risk Predictions**

It should be noted that the number of crash sites used in the model was considerably greater than the number of comparison sites. This discrepancy resulted from the separation of state highway and non-state highway sites; while the total number of rural comparison vertices chosen equalled the total number of rural crash sites the proportion of state highway sites in the comparison set was not equal to that of the crash set. This highlights a limitation of the selection method; as more crashes occurred on state highways and comparison sites were selected from regions outside the crashes' influence areas it decreased the chance of a comparison site being selected from a state highway location. This is discussed further in Section 8.3.3.



It is also important to note that the model was intended to identify potentially hazardous sites. So the comparison sites that were predicted to have high crash risks may well be dangerous. Overall, the rural state highway model suggests that roads should be designed with frequent curves that do not vary greatly in geometry (and hence operating speed) from curve to curve.

The rural state highway model was applied with the traffic flow and elevation terms held at mean values for a series of kilometre-long typical road sections; these are shown in Figure 7.2.

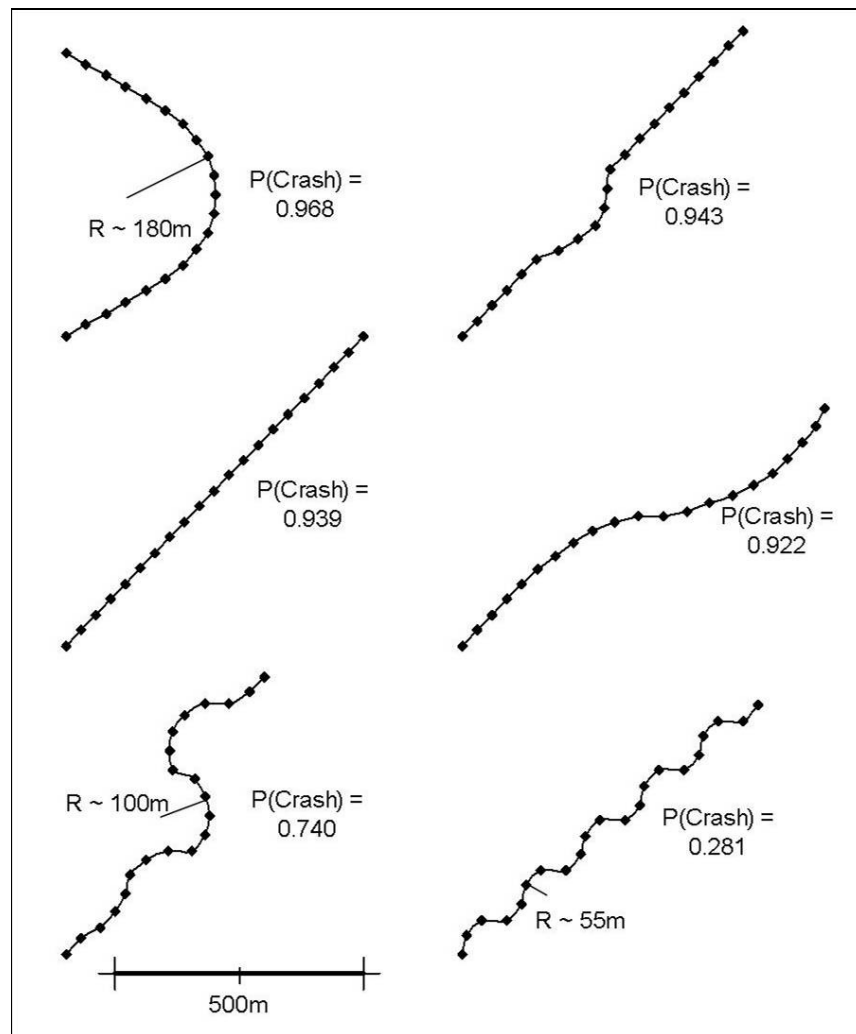


Figure 7.2 Rural State Highway Crash Probabilities

These typical sections illustrate the properties of the model. None of the sections had the combination of very low detour ratio and cumulative angle values with high standard deviation of angles shown to limit the model's accuracy.

It can be seen that the most dangerous type of road section as predicted by the model was a single large curve. Such sections have high detour ratios but relatively small cumulative angles and therefore were classed by the model as having high crash risk.

Very straight sections also appeared to have high crash risk. These sections have  $CA = 0$ ,  $SD = 0$  and  $DR = 1$  (the lowest possible values). Thus, although the detour ratio is very low it is much more influential than the cumulative angle and standard deviation of angles, which could indicate a limitation of the model.

The sections that are the most bendy are the ones that would require the slowest travelling speeds. Thus it is sensible that, for this model which is based on fatal crash occurrence, that the more bendy regions have lower crash probabilities. It is interesting to note that the section that was mainly straight with one isolated bend had a slightly higher crash risk than the completely straight section. The isolated curve probably comes as a surprise to some drivers who are expecting the road to continue the same level of demand as the straight section. This confirms the hypothesis that design consistency is important.

Sections that head in the same general direction as the straight road but have a few bends were considered slightly safer, due to increasing cumulative angle values. Sections with many small and consistent bends have similar detour ratios to the straight roads but much higher cumulative angles than individual curves so were predicted to be much safer. Thus, bendiness, when coupled with a high consistency of design, does appear to be protective.

The motivating study did not find a significant relationship between any of the bendiness measures and crash occurrence for rural state highways. The individual correlation coefficients for each of the bendiness measures were all negatively related to crash occurrence but were also not statistically significant so should not be compared with the results of this study. The motivating study did not include any

terms to represent the consistency of bendiness in the area, as this study's standard deviation of angles attempted to do. Ultimately, this study shows a more conclusive relationship between bendiness and crash occurrence and is able to confirm the motivating study's hypothesis that bendiness can be protective.

The general conclusion of studies focused on the effects of horizontal curvature on crash occurrence for isolated curves was that the occurrence of crashes increases with increasing degree of curvature (or decreasing radius of curvature). This was discussed in Section 2.1.1. However, the typical sections here show that, when added together, multiple curves with smaller radii of curvature actually have a lower crash risk than a single curve of higher radius.

#### 7.1.2 Rural non-State Highway Crash Model

The model chosen to predict the probability of crash occurrence on all other rural roads is summarised in Equation 25 and Table 7.3:

$$\log it[Y] = -1.38 + 1.19DR - 7.54CA + 0.0457SD + 0.00445MR - 0.00159E + 0.963I - 0.00632MB$$

**Equation 25 Rural non-State Highway Model**

where:  $MR$  = Mean rainfall (mm/month)

$I$  = Intersection within 30m dummy variable ( $I = 1$  if the location is within 30m of an intersection,  $I = 0$  otherwise)

$MB$  = Area of location's meshblock (km<sup>2</sup>)

**Table 7.3 Rural non-State Highway Model**

<b>Predictor</b>	<b>Coefficient</b>	<b>Standard Error of Coefficient</b>	<b>P-value</b>
<b>Constant</b>	-1.38	0.42	0.001
<b>Detour Ratio</b>	1.19	0.39	0.002
<b>Cumulative Angle (degrees/m)</b>	-7.54	0.51	<0.001
<b>Standard Deviation of Angles (degrees)</b>	0.0457	0.0073	<0.001
<b>Mean Rainfall (mm/month)</b>	0.00445	0.00130	0.001
<b>Elevation (m above sea level)</b>	-0.00159	0.00045	<0.001
<b>Intersection within 30m</b>	0.963	0.162	<0.001
<b>Meshblock Area (km<sup>2</sup>)</b>	-0.00632	0.00143	<0.001

The evaluation statistics of the model were as follows:

Test that all slopes are zero:  $G = 645$ , P-Value <0.001

Deviance goodness of fit test:  $\chi^2 = 2254$ , P-Value = 0.998

Percentage of pairs that were concordant = 82.0%

Percentage of pairs that were discordant = 17.7%

Percentage of pairs that were tied = 0.2%

**Table 7.4 Summary of Influencing Variables for Rural non-State Highway Model**

<b>Variable</b>	<b>Range</b>	<b>Mean</b>	<b>Coefficient</b>	<b>Coeff*Range</b>	<b>Coeff*Mean</b>
<b>Detour Ratio</b>	1; 5.67	1.13	1.19	1.19; 6.76	1.35
<b>Cumulative Angle (degrees/m)</b>	0; 1.77	0.280	-7.54	0.00; -13.37	-2.12
<b>Standard Deviation of Angles (degrees)</b>	0; 44.0	10.1	0.0457	0.00; 2.01	0.46
<b>Mean Rainfall (mm/month)</b>	0; 434	108	0.00445	0.00; 1.93	0.48
<b>Elevation (m above sea level)</b>	0.47; 1161	148	-0.00159	-0.00; -1.85	-0.24
<b>Intersection within 30m</b>	0; 1	0.0902	0.963	0.00; 0.96	0.09
<b>Meshblock Area (km<sup>2</sup>)</b>	0.1; 1626	40.1	-0.00632	0.00; -10.27	-0.25

The same bendiness measures (detour ratio, cumulative angle and standard deviation of angles) as for the rural state highway model were chosen. These variables had the same relationships to crash probability as in the state highway model, with cumulative angle being inversely proportional to crash probability and detour ratio being directly proportional (even though in the individual correlations it was inversely proportional). However, this time the cumulative angle measure appeared to be the most influential, which indicates that bendiness on rural non-state highway roads is very likely to be protective rather than hazardous.

Standard deviation of angles was also directly proportional to crash probability, as it was in the individual correlations, again indicating that consistency of elements is important, but less influential than bendiness considerations.

Mean rainfall was included in the rural non-state highway model. It was about as influential on the model's outcome as the standard deviation of angles variable. It was directly related to crash probability indicating that the higher the average rainfall the higher the likelihood of a crash.

Elevation was again shown to be a significant variable with higher sections of road having lower probabilities of crash occurrence. It is possible that the mean elevation term was actually a proxy measure for the true flow on links, as it was shown for the state highway roads that elevation and estimated daily traffic flow were significantly correlated.

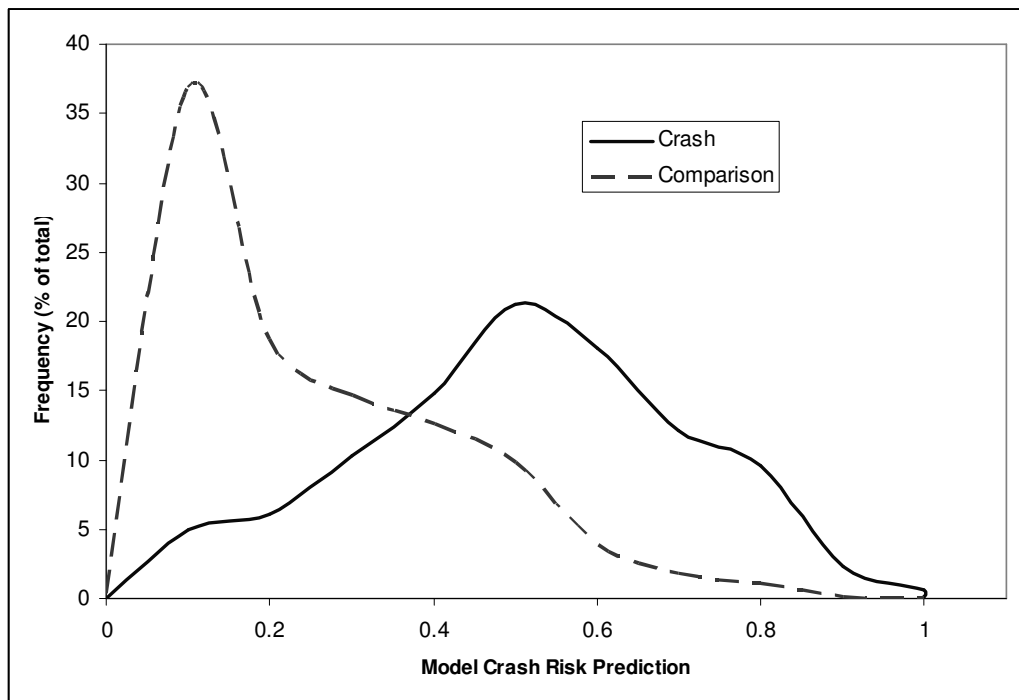
The variable that was initially intended to represent flow (number of people employed and living in a meshblock per km<sup>2</sup>) did not have a significant correlation with crash occurrence and was not useful in the model. As expected, based on correlations from the rural state highway data, it seemed that the variable chosen did not represent the flow characteristics of rural roads correctly. This may be because rural road flows are likely to be comprised mainly of through traffic and therefore numbers of people present in the area are not indicative of the total volume of traffic progressing through it. The size of the location's meshblock, however, was significant to the model. The greater the area (and hence the lower the population density, i.e. the "more rural" the area) the lower the probability of crash occurrence.

The p-values for all influencing variables included in the model were below the level of significance. Based on that, and the result of the test that all slopes were zero, it appears that the variables were appropriately chosen. However, the constant also had a very low p-value and was of similar size to the model coefficients multiplied by mean values, indicating a possibility that some important variables had been excluded from the model.

The deviance chi-squared statistic of the rural non-state highway model had a very large p-value, which indicated that the null hypothesis of the model being a good fit could not reasonably be rejected. The number of concordant pairs was very high at 82.0% which suggests that the non-state highway model was reliable.

Figure 7.3 shows the distributions of crash risk for crash and comparison sites predicted by the model. The two distributions follow the desired pattern of comparison sites at the lower end of the crash risk scale and crash sites at the higher end but the mean crash risk for the crash sites is somewhat low, which suggests that

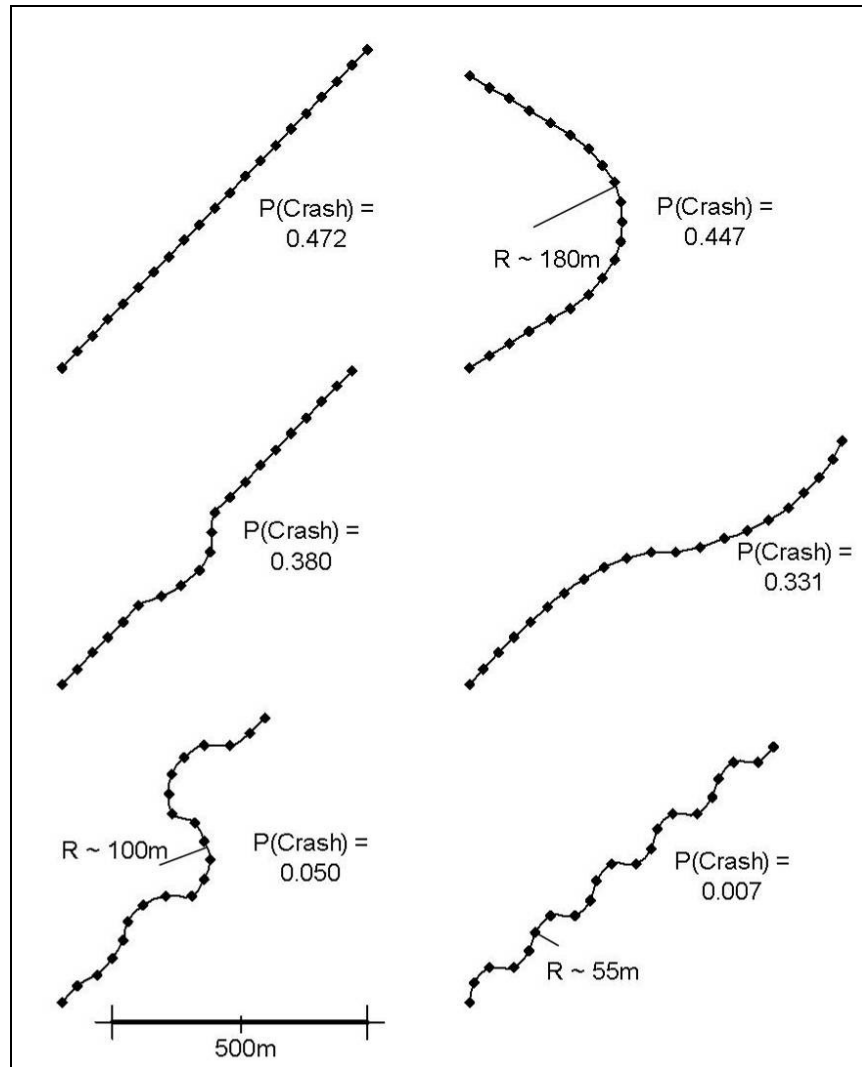
the model under-predicts crash risk. 11% of the crash sites were predicted to have a crash risk of less than 0.2; these sites generally had much higher than average detour ratio and cumulative angle values, slightly lower than average standard deviation of angles values, and were in areas of high elevation and large meshblock areas (which is assumed to signify low traffic flows). As a general rule, The rural non-state highway model should be treated with caution for sections with  $DR > 1.20$ ,  $CA > 0.37$  deg/m and  $SD < 10.0$  deg.



**Figure 7.3 Rural non-State Highway Model Crash Risk Predictions**

Again, the sample sizes of crash and comparison locations were not equal, this time with about 2.5 times more comparison data than crash data due to the selection method used.

The non-state highway model was applied to the same typical sections as for the state highway model, holding non-bendiness values at their mean levels. None of the sections satisfied all three criteria required to warrant caution of the model. The probabilities of crash predicted by the model are shown in Figure 7.4.



**Figure 7.4 Rural non-State Highway Crash Probabilities**

The probabilities give a clear indication of the relative risks of travelling on different types of rural non-state highway roads. In this case the straight road was the most risky, followed by the large single curve and then the single curve situated between two long straights. This suggests that isolated curves may have a protective effect on rural non-state highways.

The remaining four sections followed the same order as for the rural state highway model with the section of small frequent bends being the safest due to its high cumulative angle, low standard deviation of angles values and reasonably low detour ratio.



It is acknowledged that the observed trend of rural roads of higher bendiness having fewer fatal crashes than straight roads does not necessarily mean that such sections have fewer total crashes. It may simply be that slower travelling speeds on bendy roads result in higher proportions of lower-severity crashes. While further research would be required to confirm this hypothesis it should be noted that, even if this was the case, bendiness would still be protective as crashes of low severity have lower social and economic costs than fatal crashes.

Again the motivating study did not find any significant relationship between any of the bendiness measures and crash occurrence on rural non-state highway roads. It is worth noting that the detour ratio was the closest to being significant in the motivating study and was also very influential in this study. As for the state highway model, this study confirms the hypothesis of the motivating study that bendiness is protective for rural non-state highway roads.

#### 7.1.3 Urban Crash Model

The model chosen to predict crashes on all urban roads is summarised in Equation 26 and Table 7.5:

$$\log it[Y] = 0.389 - 4.83CA + 0.0858SD + 0.00436MR - 0.00236ME - 0.463JD - 0.000130PE$$

#### **Equation 26 Urban Model**

where:  $JD$  = Junction density (junctions per km)

$PE$  = Number of people residing plus number of people employed within the location's meshblock divided by the meshblock area (people/km<sup>2</sup>)

**Table 7.5 Urban Roads Model**

<b>Predictor</b>	<b>Coefficient</b>	<b>Standard Error of Coefficient</b>	<b>P-value</b>
<b>Constant</b>	0.389	0.167	0.019
<b>Cumulative Angle (degrees/m)</b>	-4.83	0.31	<0.001
<b>Standard Deviation of Angles (degrees)</b>	0.0858	0.0051	<0.001
<b>Mean Rainfall (mm/month)</b>	0.00436	0.00132	0.001
<b>Elevation (m above sea level)</b>	-0.00236	0.00046	<0.001
<b>Junction Density (junctions per km)</b>	-0.463	0.036	<0.001
<b>[Population + Employment] / Area (people/km<sup>2</sup>)</b>	0.000130	0.000024	<0.001

The evaluation statistics of the model were as follows:

Test that all slopes are zero:  $G = 906$ , P-Value <0.001

Deviance goodness of fit test:  $\chi^2 = 4327$ , P-Value <0.001

Percentage of pairs that were concordant = 77.5%

Percentage of pairs that were discordant = 22.3%

Percentage of pairs that were tied = 0.2%

In terms of choosing which of the bendiness variables to use, the bend density and cumulative angle were highly correlated so, as the cumulative angle proved to be the most effective when included in the model, the bend density variable was excluded. Similarly, the mean angle and standard deviation of angles were highly correlated; standard deviation proved to be the most effective when included in the model and added an indication of the effect of design consistency, so mean angle was excluded.

The inclusion of mean rainfall and mean elevation variables improved the model slightly so were included in the model. The standard deviation of rainfall, standard deviation of elevations and intersection variables, although each having reasonable individual correlations with crash probability, did not improve the model significantly so were excluded.

The simple junction density measure (number of junctions per km) proved to be more useful in the model than the weighted junction density measure.

The variable intended to represent flow (number of people living or working per km<sup>2</sup> of meshblock area) proved to be very significant in the model, as shown by the results in Table 7.6.

**Table 7.6 Summary of Influencing Variables for Urban Roads Model**

<b>Variable</b>	<b>Range</b>	<b>Mean</b>	<b>Coefficient</b>	<b>Coeff*Range</b>	<b>Coeff*Mean</b>
<b>Cumulative Angle (degrees/m)</b>	0.00; 4.67	0.285	-4.83	0.00; -22.57	-1.38
<b>Standard Deviation of Angles (degrees)</b>	0.0; 70.8	17.8	0.0858	0.00; 6.08	1.53
<b>Mean Rainfall (mm/month)</b>	0; 247	96.1	0.00436	0.00; 1.08	0.42
<b>Elevation (m above sea level)</b>	0; 816	61.9	-0.00236	0.00; -1.93	-0.15
<b>Junction Density (junctions per km)</b>	0.00; 6.56	2.70	-0.463	0.00; -3.04	-1.25
<b>[Population + Employment] / Area (people/km<sup>2</sup>)</b>	0; 53634	1644	0.0013	0.00; 69.72	2.14

In the urban model the probability of crash was once again negatively related to cumulative angle. The greater the angle travelled, the lower the crash risk. The

probability of crash was directly related to standard deviation of angles. The more variation between angles travelled (lower consistency), the higher the crash risk.

In terms of the other influencing variables, the probability of crash was directly related to mean rainfall. The more rain in the location, the higher the crash risk. The probability of crash was negatively related to mean elevation. The higher the location is, the lower the crash risk. The probability of crash was negatively related to junction density. The more junctions near a location the lower the crash risk. The probability of crash was directly related to the proxy flow measure. The more people living and working within an area (and hence the higher the flows on the roads) the higher the crash risk.

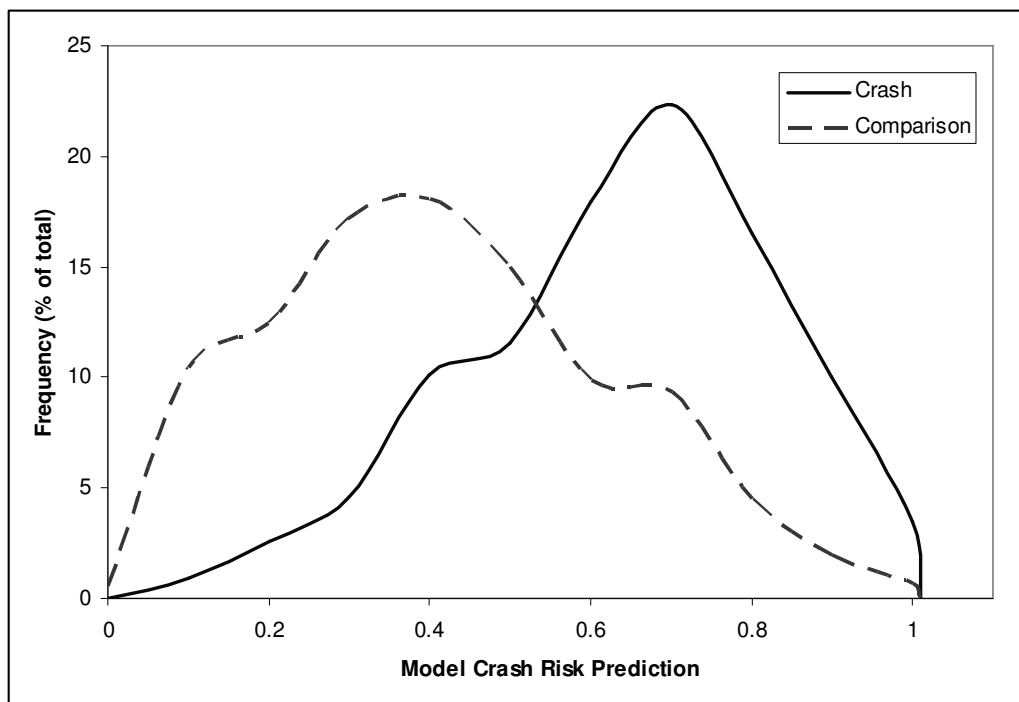
The p-values for all influencing variables included in the model were well below the level of significance which suggests that these coefficients are not zero and therefore the variables were appropriately chosen. This was reinforced by the test that all slopes are zero, which indicated that at least one of the coefficients was not equal to zero. The p-value for the constant was not above the level of significance, indicating that important influencing variables may have been excluded from the model.

The Deviance Chi-squared statistic had a p-value much less than the level of significance indicating that the null hypothesis of the model being a good fit could be reasonably rejected. However, as for the rural models, the number of concordant pairs was very high (77.5%) which suggests that the urban model was in fact reasonably reliable.

Figure 7.5 shows the distributions of crash risk predicted by the urban roads model for crash and comparison sites. The distributions follow the desired pattern, with the crash site distribution having a higher mean and very low frequency at the lower end of the crash risk scale and the comparison site distribution being the opposite. There is a considerable degree of overlap between the two distributions. 4% of the crash sites had crash risk predictions less than 0.2; these sites generally had very high cumulative angle values and very low standard deviation of angles values (i.e. many consistent bends) with other variables being reflective of the total population's mean values. 3% of the comparison sites had crash risk predictions greater than 0.8; such

sites generally had lower than average cumulative angle values and higher than average standard deviation of angles values (i.e. infrequent and inconsistent bends) but no obvious trends for the non-bendiness variables.

Thus special caution should be taken when the urban model is applied to areas of extreme bendiness ( $CA < 0.10$  deg/m or  $CA > 0.54$  deg/m) and design consistency ( $SD < 10.0$  deg or  $SD > 33.5$  deg) values. It is unlikely that the model was subject to any bias due to sample sizes as the two data sets used were of equal size.



**Figure 7.5 Urban Roads Model Crash Risk Probabilities**

The inclusion of the junction density variable and the general nature of the urban situation gives it more of a network emphasis and hence makes it harder to separate into single stretches of road as was done for the rural typical sections. Several sites chosen from Christchurch were analysed according to the urban crash probability model with mean values used for the mean rainfall, elevation and proxy-flow variables. A sample of these sites is shown in Figure 7.6.

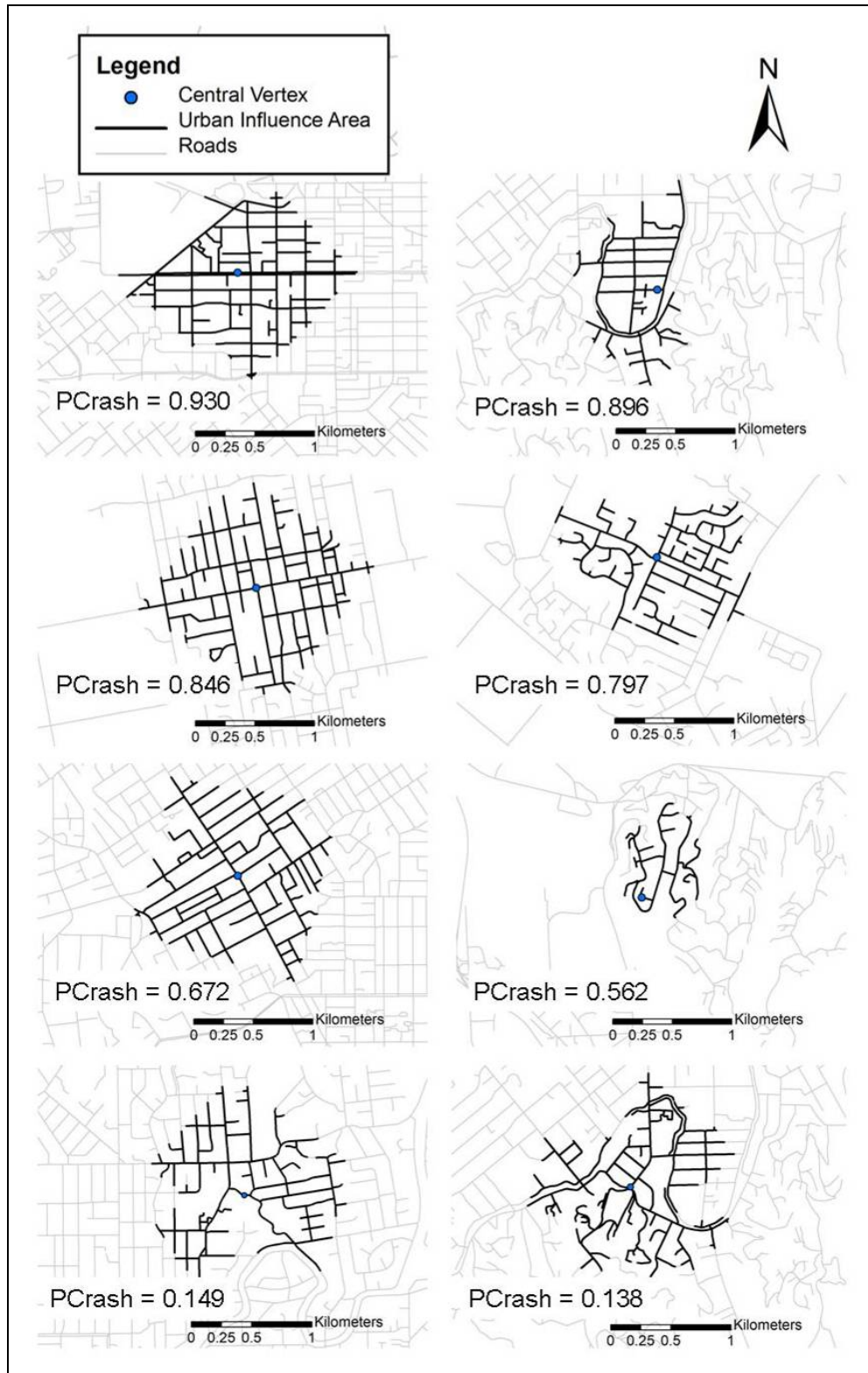


Figure 7.6 Urban Typical Examples

It can be seen from Figure 7.6 that the sites with more irregularities and higher bendiness between junctions generally have lower crash risks. These are the regions with high cumulative angle and low standard deviation of angles values (i.e. networks that are consistently bendy) and generally have lower junction densities.

It seems that the findings of the motivating study were more conclusive than those of this study for urban roads. Whereas the model developed here did not have a statistically significant fit, the motivating study found significant negative relationships for two of the bendiness measures; cumulative angle (which also appeared in this model) and detour ratio.

It seems though that this study, which produced significant results for the two rural cases, has highlighted the inappropriateness of applying the Fowler method in the urban situation. It is much harder to infer the route taken by a driver involved in an urban crash due to the complexity of the urban networks and there are many other complicating factors associated with higher traffic flows, increased conflict points and generally shorter trips. While the aim of this study has not been to suggest an alternative to bendiness as a predictor of crashes on urban roads, it has shown that a method applicable to rural roads should not be used on urban roads.

#### 7.1.4 Comparative Analysis of Models

The main difference between the two rural road models was that the state highway model included actual flow data. The non-state highway model attempted to account for exposure by including the meshblock area as an influencing variable but it is doubtful that this variable was an appropriate substitute. While the meshblock area had the highest correlation with AADT of all variables tested for the state highway data the correlation was still reasonably low (-0.281). Mean elevation and mean rainfall also had similar correlations with AADT and it is likely that these measures also appeared in proxy for traffic flow.

If, however, the mean rainfall term was not entirely a proxy measure its relationship suggests that non-state highways are more susceptible to variations in weather. This may be an indication of the different standards of the two road types; state highways having typically larger flows are generally provided with better drainage, more

frequent maintenance and better treatments applied to prevent negative effects that result from rainfall.

The probabilities for crashes on non-state highway rural roads (Figure 7.4) were much lower than for state highways (Figure 7.2) of the same geometries. The lower probabilities do not necessarily mean that rural non-state highway roads are safer than state highways as state highways have greater traffic flows and hence greater exposure. Also, the models should not be compared quantitatively with each other as they involve different variables and the rural non-state highway model did not take into account actual flow values.

At the higher end of the crash risk scale, the two rural models did not produce the same ranking of road section types. The state highway model suggested that isolated curves were more dangerous than straight sections but the non-state highway model gave the opposite result. This may be due to the different properties of state highway and non-state highway roads; drivers may be more cautious when approaching curves on non-state highway roads as they do not expect as high a level of design as for state highways. On the other hand, the different rankings may be due to the different data available to the two models; in which case the state highway model is more reliable.

It appears that the method was generally successful when applied to both rural situations. The consensus between the two rural models was that bendiness, accompanied with high design consistency, is protective against crashes on rural roads.

The urban model had several differences to the rural models. The measure intended as a proxy for flow was a very influential component of the urban model but not used in the rural models. This is consistent with the idea that rural flows are comprised mainly of through traffic whereas, for the urban situation, the vehicle flows are directly related to the number of people living and working within the immediate area. While this is probably still not the best way of differentiating between flows in similar locations (e.g. two central Auckland intersections) it should show a good distinction between locations of greater variation (e.g. a central Auckland intersection and an



intersection in Gore). Thus, when applied to data for the entire country, the urban model should account well for variations in exposure to the possibility of crashes.

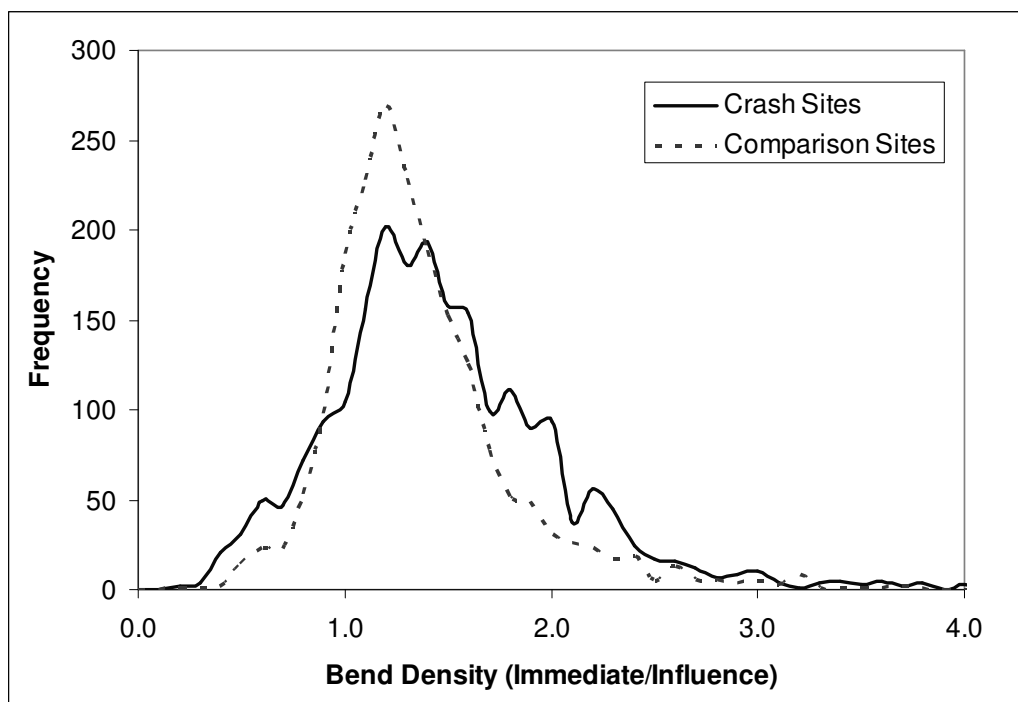
The inclusion of the junction density variable in the urban model (which was not present in the rural models) suggests that the presence of intersections is more important in the urban situation than the rural. This is sensible given that intersections are much more closely spaced in urban areas. In this situation the junction density is effectively another bendiness variable as junctions involve sharp angles for turning traffic. The junction density variable had a similar influence on the model to that of the other two bendiness measures.

However, when not coupled with an indication of the flow along each link, use of the junction density variable may result in misleading results in areas with structured road hierarchies. For example, the presence of many cul-de-sacs would increase the junction density value but very little traffic is likely to travel there in reality. The fact that fatal crash risk decreases with increasing junction density may be an indication of the slower speeds (and hence lower crash severities) that result in urban areas with high traffic volumes and many intersections. Locations with fewer junctions enable vehicles to travel faster and may be specifically designed with higher speed limits (e.g. major arterial urban roads).

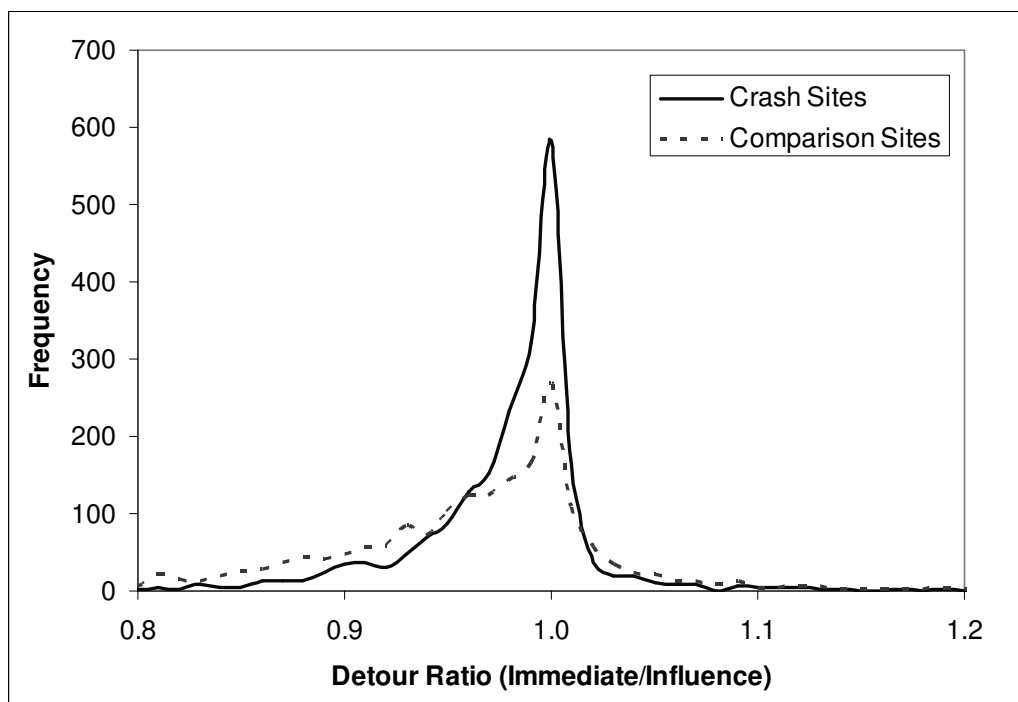
Overall, the urban model was not as effective as the rural ones and it seemed that the method was not as appropriate for the urban situation.

## **7.2 Immediate versus Influence Area Bendiness Measure Evaluations**

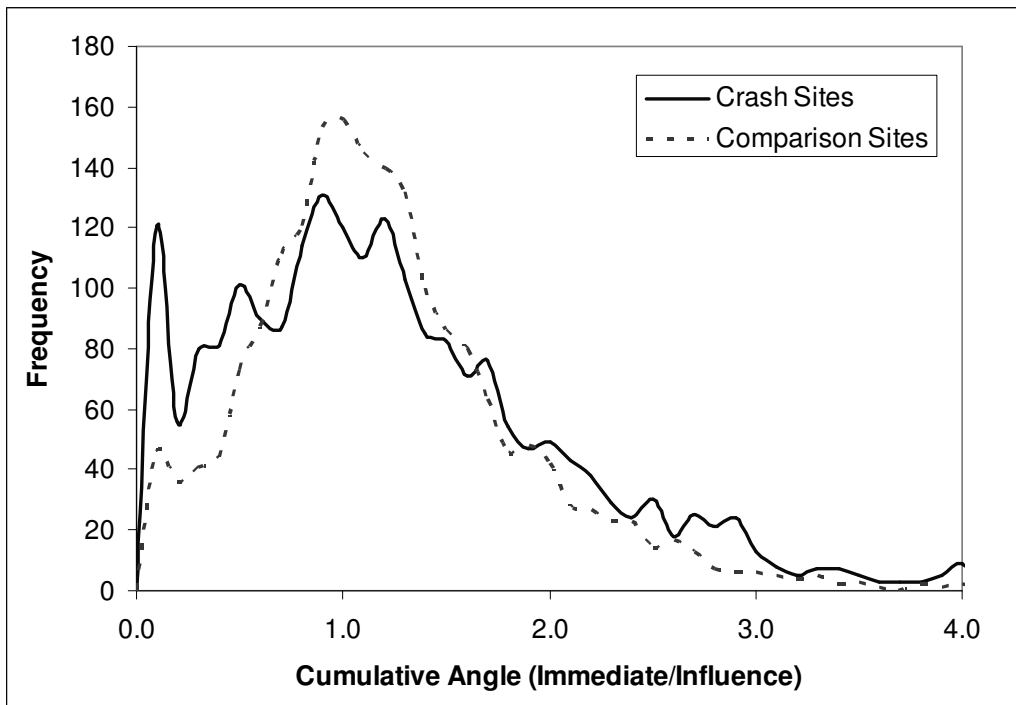
The design consistency evaluation method that compared bendiness measures of the immediate area (within 250m of the location of interest) and influence area (within 1km) was applied to all rural crash and comparison sites according to the method detailed in Section 5.3.2. Results for this investigation are shown in Figures 7.7 to 7.11 and Table 7.7.



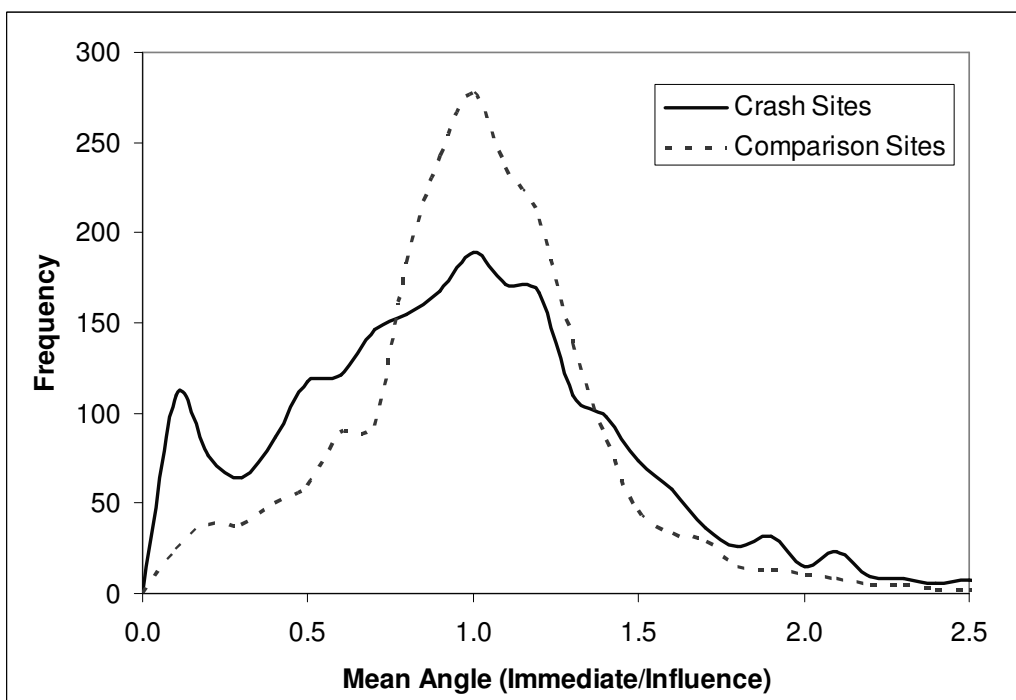
**Figure 7.7 Immediate versus Influence Area Bend Density for Rural Sites**



**Figure 7.8 Immediate versus Influence Area Detour Ratio for Rural Sites**



**Figure 7.9 Immediate versus Influence Area Cumulative Angle for Rural Sites**



**Figure 7.10 Immediate versus Influence Area Mean Angle for Rural Sites**

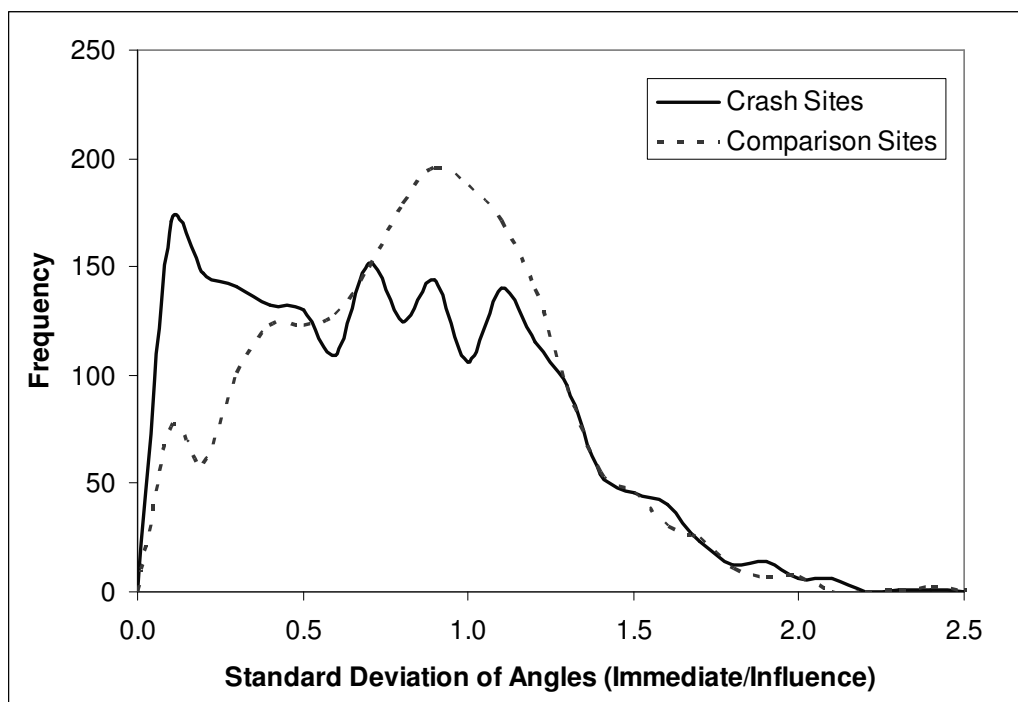


Figure 7.11 Immediate versus Influence Area Standard Deviation of Angles for Rural Sites

Table 7.7 Immediate vs. Influence Ratio Statistics

	Immediate versus Influence Area Ratio for:				
	Bend Density	Detour Ratio	Cumulative Angle	Mean Angle	Standard Deviation of Angles
<b>Crash Site Average</b>	1.423	0.967	1.199	0.931	0.727
<b>Crash Site Standard Deviation</b>	0.558	0.075	0.837	0.557	0.528
<b>Comparison Site Average</b>	1.317	0.951	1.163	0.966	0.800
<b>Comparison Site Standard Deviation</b>	0.446	0.125	0.645	0.430	0.411

An immediate/influence ratio of one indicates that the location of interest is a typical example of its surrounding area. For all five measures, both the comparison sites and crash sites have very similar mean immediate/influence ratios that are close to one. However, except for the detour ratio result, the mean values of the comparison sites are all closer to one.

A smaller standard deviation means that individual values are closer to the mean value. Again with the exception of the detour ratio result, the comparison sites have less spread.

These two observations indicate that the non-crash locations have more consistent geometries as they have mean values closer to one and lower standard deviations than the crash locations. This suggests that increasing the consistency of elements with respect to their surrounding locations reduces the likelihood of crash occurrence.

This method uses the same bendiness measures as the binary logistic modelling approach but in a different way; parameters are normalised by comparing measures over different lengths and thus there is no need for inclusion of other influencing factors. If any of the measures gave a clear indication of the differences between the crash and comparison site properties this method would be advantageous over the modelling method as it would only require the computation of one variable.

However none of the measures appear significantly different between the crash and comparison sites as the distributions all peak at similar points. Thus, by this analysis it would be difficult to prescribe a certain range for any of the bendiness measures that would give an adequate probability of no crash occurrence.

Therefore, it seems that while the method of comparing immediate and influence area characteristics for a location can have some use in comparing crash and non-crash sites it would be preferable to use the binary logistic model developed in the identification of potential crash sites.

## **8.0 Conclusions**

This section summarises the findings of this study and their implications on the study objectives. A list of problems encountered is given in hope that future researchers will be able to learn from the challenges faced in the course of this study. Finally, a series of suggestions for further ways of researching the effect of road network bendiness on traffic crash occurrence is given.

### **8.1 Summary of Findings**

#### 8.1.1 Initial Comparisons with Motivating Study's Measures

When the initial results of this study were aggregated at a TLA level and compared to the results of the motivating study, the observations varied. The bend density measure was found to be the most consistent between the two studies, indicating that the number of bends per kilometre is generally consistent throughout a TLA. However, as this measure was not included in any of the models developed in this study its consistency with the motivating study is not very focal. It was shown that the bend density and cumulative angle measures were closely related (as they are both affected in the same way by changes in resolution of data) and although cumulative angle was the favoured measure it would be feasible to replace it with bend density.

The cumulative angle measure, which was used in all three of the models developed in this study, had reasonably consistent values when compared with results of the motivating study. Results for the detour ratio, another important bendiness measure in the models of this study, did not appear to be consistent with those of the motivating study when compared by scatter plot. The detour ratio comparisons did, however, have the lowest percentage errors. This was attributed to slightly different definitions and very different applications between the two studies.

The initial justifications of undertaking this study were concerned with aspects of the motivating study's methodology. Had the initial results of this study been completely consistent with those of the motivating study, it would indicate that this study's approach was no more appropriate than the aggregated approach of the motivating study, as bendiness values for a whole TLA were representative of the bendiness values of individual crash sites. Overall, it was determined from the initial

comparisons that there was enough difference between the results of the two studies to warrant continuation of the proposed research method.

#### 8.1.2 Crash versus Non-Crash Comparisons of Bendiness Measures

It was determined that findings of the preliminary comparisons between bendiness measures for crash and non-crash locations were not very reliable as they did not take into account the differences in traffic flow between the sites. The rural non-state highway sites (probably the sample with the highest flow variability) was the least consistent between crash and comparison site properties.

This exercise gave a useful indication of the distributions of the bendiness measures that were used later on in the modeling process. However its main conclusion was that, unless important non-bendiness variables (such as traffic flow) were kept consistent between crash and comparison sites, further investigation would be required to make reliable conclusions.

#### 8.1.3 Use of Flow Weighting Techniques for Bendiness Measure Calculation

A major concern of this study was that the lack of actual flow data available for weighting of possible routes would result in large inaccuracies. To test the effects of weighting routes according to flow along links a sample of flow data were obtained for the Upper Hutt City TLA.

The majority of locations had low flow variability and consistent results for the two weighting methods. Obvious differences, however, between the flow-weighted and non-flow weighted results occurred for areas of high flow variability. Thus, the binary logistic regression models developed without flow weighting should be applied with caution. It is likely that they will not be as accurate in locations with clearly defined road hierarchies.

The Upper Hutt City TLA did not contain a large enough crash population to develop new binary logistic regression models with bendiness measures weighted by flow. This investigation should however serve to highlight the limitations of the models developed and illustrate how a better model could be developed, if flow data were available.

#### 8.1.4 Binary Logistic Regression Models

Binary logistic regression models were developed for the rural state highway, rural non-state highway and urban cases to compare the effects of bendiness for crash and non-crash sites while accounting for other, non-bendiness related variables.

The model developed for crash risk on rural state highways achieved a significant goodness-of-fit test result and high number of concordant pairs, which indicated that it was a suitable model. This model showed large individual curves to be the most hazardous section types, followed by single curves situated in otherwise straight sections. The section of road tested with this model that was shown to be the safest was a series of small and consistent bends. An advantage of the rural state highway model over the other two models was that it was based on actual flow data.

The model developed for crash risk on rural non-state highway roads also achieved a significant goodness-of-fit test result and high number of concordant pairs. It was slightly disadvantaged compared to the rural state highway model, as there was no actual flow data available. The census meshblock area (which is defined based on number of residents) was shown to be the most appropriate proxy for flow. The rural non-state highway model predicted straight sections of road to be the most dangerous, followed closely by large individual curves. As for the state highway model the section of road with many small and consistent bends was shown to be the safest. It was concluded that the two models' results should not be compared quantitatively due to the difference in flow measures used.

The model developed for crash risk on urban roads did not have a statistically significant goodness-of-fit result, indicating that the method developed in this study may not be applicable to the urban situation. The model did however produce a reasonably high number of concordant pairs and so was still applied to typical sections of urban network. It appeared that consistently bendy areas with low junction densities were the safest type of urban area.

The results of these three models were very different to those of the motivating study. Whereas this study showed significant trends between bendiness and crash occurrence



in rural situations, the motivating study found no such evidence. Conversely, the motivating study found a significant relationship between bendiness and crashes on urban roads whereas this study's urban model was not statistically significant. It was assumed from the start that bendiness may not be a major contributor to urban crash occurrence and, contrary to the results of the motivating study, the models developed here suggest this is true.

Given that this study returned such different results to those of the motivating study it would not be appropriate to compare the results of this study with those of Haynes et al's (2007a) British study as was initially intended.

#### 8.1.5 Immediate versus Influence Area Analysis

An attempt was made to normalise the bendiness measures so that they could be applied to locations without requiring information on other variables. This was done by calculating the bendiness measures over the immediate area (within 250m of the crash or comparison location of interest) and comparing with the bendiness measures over the influence area (within 1km of the crash or comparison location).

It was found that the non-crash locations used as comparison sites were generally more consistent (i.e. their distributions had a lower spread) than the crash locations. However, the two location types generally had similar mean immediate/influence ratios. Therefore, it would be difficult to use this technique to identify problem sites as a large proportion of the crash sites still fell within the non-crash site tolerance.

Thus, although this method was simpler and required less data and time to calculate, it would be preferable to include other variables and combine bendiness measures using the binary logistic regression technique.

### **8.2 Implications of Findings with respect to Objectives**

The first objective of this research was to improve the method of analysis used in the motivating study by providing alternative measures of bendiness, considering the importance of influencing variables and taking into account factors such as flow choices at intersections.

A new measure was employed, the standard deviation of angles measure, which was used to represent design consistency and proved to be influential in all three of the binary logistic regression models developed. The bendiness measures used in the motivating study were determined to be useful in principle and the focus of this research was finding the most appropriate method of applying them. The measures were applied to influence areas of crashes and comparison sites, rather than at an aggregated level as for the motivating study. This required a new method of analysis and the binary logistic regression model was chosen.

In addition to the modelling method, the immediate/influence area method was developed but, as detailed previously, it was decided that the modelling method was preferable. Thus, the first objective was satisfied as a new measure and two new methods were investigated.

The second objective of this research was to determine the relationship between network bendiness and traffic crashes in New Zealand using the revised measures and methods. Significant relationships were found between road bendiness and crash occurrence for the two rural road cases thus clearly satisfying this objective. While the rural cases analysed all possible routes from a location to the extent of its influence area, the exclusion of any variable representing junction density suggests that the “network” part of this objective was not very important. As the definition of bendiness implies that many elements are analysed it would be sufficient to say that the effect of road bendiness on traffic crashes has been analysed.

The urban model, with its inclusion of a junction density term, gives a better example of how network effects were considered. However, the urban model did not have a significant goodness-of-fit. The lack of significance of the urban model does not necessarily indicate that the objective was not met, rather it is probably more an indication that there were other predictors of urban crash occurrence more important than bendiness in the urban network.

The third objective was to compare and contrast results of this study with those of the motivating study and Haynes *et al*'s (2007a) British study. It should be noted that this

objective was made before the conclusions of the motivating study were reached. It is difficult to compare the two studies given that the only significant result for the motivating study came from the urban situation but the urban model developed here did not have a significant goodness-of-fit and it was considered the method would not be applicable to the urban situation.

The fact that the results were so different from those of the motivating study is in itself a very useful comparison. The method of application used in this study was very different than that of the motivating study and its significance in rural applications suggests that its localised approach is much more appropriate than an aggregated method.

Finally, the fourth objective of this research was to determine the relevance of findings to New Zealand practices for road design. This objective is catered for in the analysis of typical road sections. It can be seen that, for rural roads, sections with frequent, small and consistent bends are much safer than straight roads. Thus road designers should consider the option of introducing bends to roads that could actually be straight in order to keep the level of demand imposed on drivers at an optimum. Consistency has also been shown to have an important effect and elements should be considered in the context of the wider road environment.

Obviously there are other factors involved in road design; if not limited by topography a bendy road would seem less economically feasible than a straight one. However as this research implies that crash risk decreases with increasing bendiness and since the economic impact of a fatal crash is very high, when all costs are considered building bendy roads may be economically more feasible.

### **8.3 Problems Encountered**

#### **8.3.1 Computer software**

As mentioned in Section 4.2.4, different methods and computer packages were trialled to find the most appropriate technique of modifying the road network data. The main reason for not initially using ArcMap was that the university's Engineering Department does not have the ArcInfo extension license that contains some of the

commands required. Eventually the student chose to use the license held by the university's Geography Department.

This in no way suggests that the Engineering Department is poorly equipped; in general there is little need amongst engineering students for this extension. However, this situation does highlight the effect software can have on the analysis process; it took two weeks to process the buffer area around one crash by using CAD software and manual operations, but this can be done for the whole NZ road network with ArcInfo in a matter of days.

### 8.3.2 Data formats

Much of this research involved processing data (mainly the road network, but also crash and flow) into formats applicable to the investigation. Ideally a road network for the whole of New Zealand with comprehensive flow information and no duplicate links would have been used. Crash data directly corresponding to the road network, with route information would also be ideal. This would give more accurate results and allow more time to be spent on analysis rather than preparation. Data collection is expensive but perhaps these findings can help steer the decisions as to what data is collected and in what format in the future.

### 8.3.3 Selection of comparison sites

It was acknowledged that the unequal sizes of the rural crash and comparison site data sets (when sub-divided into state highway and non-state highway) may have disadvantaged the rural models. This problem could have been prevented by separating the state highway and non-state highway sites from the start of the analysis and then selecting appropriate numbers of comparison sites.

The fact that the ratio of state highway to non-state highway comparison sites was much lower for the crash sites illustrates another problem faced in this research. It was determined that comparison sites should be selected from parts of the network that were not contained by the influence areas of the crash sites. As a large proportion of the crash sites were located on state highways this meant that there was less chance of state highway vertices being selected as comparisons. This would not have mattered as much if comprehensive flow data were available for the whole road

network but as it were not it would have been better if the exposure characteristics of the comparison sites were chosen to reflect those of the crash sites.

This signifies the importance of preparation before making calculations. While there will always be some unforeseen circumstances encountered along the course of research it is important to have a good understanding of the requirements of the data to be used so that it can be properly processed before starting the calculation phase. While the requirements were generally identified well in this study the problems encountered with the selection of comparison sites should serve as a warning for researches undertaking similar investigations in the future.

## **8.4 Areas for Future Research**

### **8.4.1 Investigation of influence area sizes**

For this thesis an influence area of 1km was used exclusively. It was anticipated that this value, although loosely based on previous research, would require further investigation. However, given the time taken to perform the computations for the whole of the country, it was not possible for this research to investigate the optimum size of influence areas used.

The evaluation of bendiness measure ratios for immediate versus influence areas that was intended to represent design consistency may also provide useful information for determining the optimum influence area size. The results from this investigation showed that the bendiness of the 250m immediate area was generally reflective of the bendiness of the 1km influence area. This may indicate that a strict definition of influence area size is not necessary.

More research should be done into the psychology of crash occurrence so that the time of travel that actually contributes to driver's impressions of the crash site is understood. Based on this, a further investigation, perhaps focusing on a subsection of the data used in this research, could be performed to determine the optimum influence area size. It is suggested that the sizes be based primarily on travel time rather than distance and therefore area covered would be determined by vehicle speeds within the area. The first major obstacle in this would be finding sufficient

data regarding vehicle speeds. A basic method would be to assign estimated vehicle speeds based on whether or not the location was in a rural or urban area; this may not be very advantageous over the 1km influence area method which created different models for rural and urban areas.

Within rural and urban areas there is still a lot of variation in speed limits. Therefore, a more advanced method of estimating vehicle speeds would require speed limit data for each link. Even so, speed limits are not always appropriate indications of the actual speeds taken by vehicles due to congestion effects and drivers who disobey speed limits. Therefore, an even better data set would contain speeds based on actual observations of vehicle speeds, such as that of the traffic flow model provided for the Upper Hutt City Council region. This study found that such data were not readily (or cheaply) attainable for regions large enough to gain statistically significant results.

Once the bendiness measures for a variety of different influence area sizes had been computed some form of evaluation method would be required to find which was the most suited to determining crash occurrence. This could be done similarly to the binary logistic regression modelling done for the 1km influence area with the model statistics then compared to find the influence area that resulted in the best model performance.

#### 8.4.2 Alternative weighting methods

When re-considering the most appropriate influence area to use it may be beneficial to adjust the methods of weighting alternative routes. The main premise of the influence area method (and in fact this entire study) was that the roads travelled for a certain distance previous to a crash location had affected the driver's perceptions and directly attributed to the crash occurrence. However, it was also assumed that the true distance was not known and that different influence areas should be tested to find the most appropriate distance to be used. An investigation into influence areas may yield a suggested value of influence area size. It is, however, unlikely that this value would represent some unseen threshold that divides road sections of equal influence on crash outcome from sections with absolutely no influence.

It would be more appropriate to assume that the further away a section of road was from the crash site, the lower it's influence on the outcome of the crash was. The size of the influence area would therefore correspond to the point at which the road sections were so far away from the crash site they were deemed have an insignificant effect on the outcome.

Such a weighting method should be incorporated with those previously suggested, preferably those that take into account the flow along links. An example is given below of how a distance term,  $T_l$ , could be added to the equation from Method 4:

$$BM_{Crashj} = \frac{\sum_{r=1}^R \left[ \frac{BM_{Routejr}}{\sum_{l=1}^L \frac{d_l T_l}{f_l}} \right]}{\sum_{r=1}^R \left[ \sum_{l=1}^L \frac{d_l T_l}{f_l} \right]^{-1}}$$

**Equation 27 Weighting Method 4 Modified**

Where:  $BM$  = type of bendiness measure (detour ratio, mean angle etc)

$Crash_j$  = one particular crash

$Route_{jr}$  = a route belonging to the influence area of crash<sub>j</sub>

$R$  = the total number of  $route_{jr}$ s in the influence area.

$f_l$  = the flow along link  $l$

$d_l$  = the length of link  $l$

$T_l$  = the travel distance from the start of link  $l$  to the crash location

$L$  = the total number of links along a route

#### 8.4.3 Determination of actual routes travelled

A major pitfall of this study was the inability to distinguish the actual route travelled by drivers previous to crashing from all possible routes in the influence area. Attempts were made to predict the most likely routes taken and weight all possible routes according to traffic flows, however exposure may not be the biggest crash influence and in many cases it may have been the route least travelled that was the most hazardous. Therefore, more research into determining the route travelled would be very beneficial to this research and most probably any other road safety investigations.

Use of a single route would also allow much faster computations per crash and hence reduce the computing time of the study. This would enable more analysis to be done in the same amount of time and other methods to be tested.

As discussed previously there were several obstacles to determining routes taken based on CAS information, mainly the lack of information given and the ambiguity of data fields used. Further research may only be able to suggest a new method of crash data collection that could be implemented in future to allow subsequent studies to properly assess routes travelled.

#### 8.4.4 Incorporate the effects of hilliness

Several studies detailed in Section 2.1 used hilliness as a variable. Bjorketun (2005) found that, for Swedish high-speed rural locations, hilliness was a more significant contribution to crash occurrence than bendiness.

This study, in accordance with the motivating study, accounted for variations in road height by using the average and standard deviation of road height as variables in the modelling process. This could be improved by calculating hilliness measures according to the same principles of the bendiness measures thus giving a more precise indication of the effect of the vertical component of road alignment on accident occurrence.

#### 8.4.5 Include injury crashes in analysis

This research only analysed the effect of New Zealand's road network bendiness on fatal traffic crashes, rather than including injury or even non-injury crashes. This was done primarily to allow comparison with the motivating study, which considered that poor reporting rates of non-fatal crashes and high variations of reporting rates between TLAs would limit the study's accuracy. It was also recognised in retrospect that to analyse all injury crashes reported over the past ten years would have huge demands in terms of computing power and time, given that there were approximately 40 times more injury crashes than fatal crashes and the fatal crashes alone took a long time to analyse.



Inclusion of injury crashes in the analysis would, however, bring some benefits to the study. Having a larger sample size may counter the effects of under-reporting and produce a more statistically significant model. Alternatively, it was recognised that many changes had been made to the road network over the ten-year period and parts of the digital network did not necessarily well represent the roads at the time of crash. Thus including injury accidents could enable the use of a smaller time period for a similar sample size.

Future research should be done to compare models for the effects of road network bendiness on injury crash occurrence with those created by this study for fatal crash occurrence. This may also allow a better understanding of the effects of road bendiness on crash severity. For example, one hypothesis may be that although fewer fatal crashes occur on bendy roads this is simply due to lower travel speeds and hence lower severity when crashes do occur.

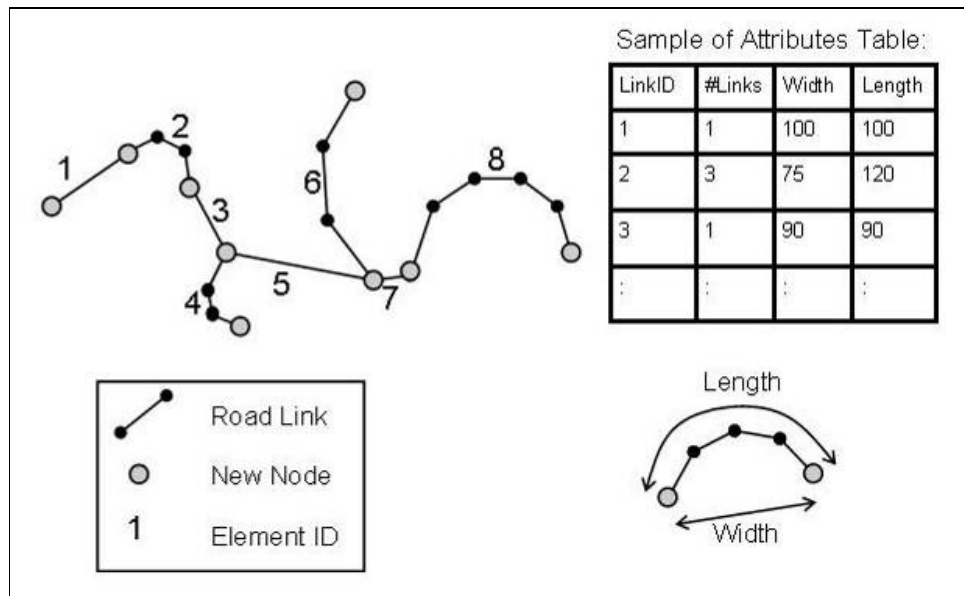
#### 8.4.6 Design alternative measures to represent design consistency

With the exception of the standard deviation of angles term all of the measures used in this study were basically those of the motivating study applied in different ways. While these measures captured the important aspects of bendiness and were consistent with previous studies it is envisaged that future research could develop new measures, especially some more suited to GIS applications.

One way of doing this might be to pre-process the road network before analysing it with respect to crashes. After splitting arcs into individual links the links could then be regrouped according to the elements (curves or tangents) they belonged to. Each element could be gauged in terms of its length and total curvature or other measures such as inferred radius of curvature etc. Thus, instead of sorting through all the individual links of a route and computing bendiness measures the characteristics of the elements forming the route could simply be aggregated.

Pre-processing the road network would mean that individual calculations of bendiness measures around crash locations would be faster as elements would be more aggregated and the effects of having to repeat calculations for similar locations would be minimised. Figure 8.1 shows an example of where a road network has been

processed so that the links belonging to each element have been grouped together and new nodes created at element ends. Each element has had a width and length value calculated; these should be used to compare different curves. Obviously this is just an example of what could be done and other fields would be included.



**Figure 8.1 Pre-processed Road Network Concept**

Some preliminary attempts of this method were made in conjunction with this study but have not been discussed in this thesis as the technique was not improved to a sufficient standard to gain meaningful results. The biggest problem encountered was how to sort through the network and identify which links belonged to which elements.

#### 8.4.7 Relate findings to knowledge of rate of crashes per individual curve

This was one of the initial objectives of the study, made before much investigation had been performed, but was omitted from the revised objectives. It was not achieved as the formulation and application of the new method required much more time and resources than initially anticipated and also because other avenues, such as the application of the measures in the immediate/influence area method were explored.

However, this would still be a useful investigation to make as finding a link between bendiness and individual curves would enable the data from studies of isolated curves to be used in the study of bendiness thus significantly increasing the sample size.

Such an investigation should focus on interpreting the bendiness measures in a more traditional way, by expressing them as curve radii or design speeds etc and tying this in with previous studies of horizontal curvature that related crash occurrence to curve radius.

## **9.0 Acknowledgements**

There are many people I wish to thank for their input, advice and support throughout the completion of my thesis.

I am very grateful to Robin Haynes (School of Environmental Sciences, University of East Anglia, UK) whose original studies provided me the motivation for this research and who was continually helpful in providing advice, clarification and much-needed data. Also, thanks to all the other members of the motivating studies, especially Iain Lake (School of Environmental Sciences, University of East Anglia, UK) who kindly and patiently answered many difficult GIS-related questions by email.

Clive Sabel, another member of the motivating study at the time based at the University of Canterbury's Geography department, was very helpful throughout the process, especially in giving advice on the method to be applied and GIS techniques. Simon Kingham and Jamie Pearce (also from the Geography department) were also very helpful in providing data and explanations of the motivating study.

My understanding of GIS techniques was also greatly enhanced by help from John Thyne, the Geography department's GIS manager who, although neither directly associated with the motivating study or my university courses, was always willing to take time to explain concepts, make suggestions and help tackle difficult problems. I am extremely grateful for John's help. Many thanks are due also to Brandon Hutchison, the Civil Engineering department's computer analyst who provided me with a lot of IT support and advice.

I am extremely appreciative of the efforts made by Fergus Tate and Jason Wildman (MWH Wellington) to provide me with the Upper Hutt City Council flow data. Also to Manu King (Transit New Zealand) who provided the state highway spatial flow data at very short notice with much encouragement and support.

Alan Nicholson, head of the University of Canterbury's civil engineering department, has been a wealth of information and inspiration throughout my studies and I am especially thankful for him always taking time out from his busy schedule to help me.

Andre Dantas, my co-supervisor, has provided a lot of useful feedback on my methodology and report structure. I am very grateful to Andre for all his support, for demonstrating his passion for transportation engineering and his care and concern for his students.

Finally I wish to thank my primary supervisor, Glen Koorey, who has been a major influence on my choice of study and was the one to suggest this thesis topic. Glen has been a great motivator, and source of direction and inspiration. I am extremely indebted to him for all he has taught me, especially over the last two years.

## 10.0 References

- AASHTO (American Association of State Highway Transportation Officials). (1990). *A Policy on the Geometric Design of Highways and Streets*, AASHTO, Washington D.C.
- AASHTO. (2001). *A Policy on the Geometric Design of Streets and Highways*, AASHTO, Washington D.C.
- Abdel-Aty, M. A., and Radwan, A. E. (2000). "Modeling traffic accident occurrence and involvement." *Accident Analysis and Prevention*, 32(5), 633-642.
- ACC (Accident Compensation Corporation), and LTSA (Land Transport Safety Authority). (2000). "Down with Speed: A Review of the Literature and the Impact of Speed on New Zealanders."
- Agresti, A. (1996). *An Introduction to Categorical Data Analysis*, John Wiley and Sons, New York.
- Agresti, A., and Finlay, B. (1997). *Statistical Methods for the Social Sciences*, Prentice Hall, New Jersey.
- Alsop, J., and Langley, J. (2001). "Under-reporting of motor vehicle traffic crash victims in New Zealand." *Accident Analysis & Prevention*, 33(3), 353-359.
- Appleton, I., Hannah, J., Noone, M., and Wilkie, S. "Road Infrastructure Safety Assessment." *22nd ARRB Conference - Research into Practice*, Canberra, Australia.
- Austrroads. (2003). *Rural Road Design: A Guide to the Geometric Design of Rural Roads*, Austrroads Incorporated, Sydney.
- Bjorketun, U. "Accident Data and Road Alignment Measures." *3rd International Symposium on Highway Geometric Design*, Chicago.
- Cairney, P. (2005). "Further Aspects of the Relationships between Geometric Alignment and Crashes." *AP – R279/05* Austrroads Inc., Sydney.
- Castro, M., Pardillo-Mayora, J. M., and Sanchez, J. F. "Alignment Indices as a Tool to Evaluate Safety and Design Consistency in Two Lanes Rural Roads." *3rd International Symposium on Highway Geometric Design*, Chicago.
- Charlton, S. G. (2007). "The role of attention in horizontal curves: A comparison of advance warning, delineation, and road marking treatments." *Accident Analysis & Prevention*, In Press, Corrected Proof.
- Drakopoulos, A., and Ornek, E. (2000). "Use of Vehicle-Collected Data to Calculate Existing Roadway Geometry." *Journal of Transportation Engineering*, 126(2), 154-160.
- Echaveguren, T., Bustos, M., and de Solminihac, H. (2005). "Assessment of horizontal curves of an existing road using reliability concepts." *Canadian Journal of Civil Engineering*, 32(6), 1030.
- Elvik, R. (2006). "Laws of accident causation." *Accident Analysis & Prevention*, 38(4), 742-747.
- Elvik, R., and Vaa, T. (2004). *The Handbook of Road Safety Measures*, Elsevier, Amsterdam.
- ESRI. (2006). "ArcMap 9.2." Redlands, California.
- Figueroa, A. M., and Tarko, A. P. "Free-flow Speed Changes in the Vicinity of Horizontal Curves." *3rd International Symposium on Highway Geometric Design*, Chicago.
- Gibreel, G. M., Easa, S. M., Hassan, Y., and El-Dimeery, I. A. (1999). "State of the Art of Highway Geometric Design Consistency." *Journal of Transportation Engineering*, 125(4), 305-313.

- Google Earth. (2007). "Map of Port Hills Region." available:  
["http://earth.google.com/".](http://earth.google.com/)
- Grime, G. (1987). *Handbook of Road Safety Research*, Butterworth & Co., Cornwall.
- Grubestic, T. H. (2007). "Zip codes and spatial analysis: Problems and prospects." *Socio-Economic Planning Sciences*, In Press, Corrected Proof.
- Harris, R., and Johnston, R. (2003). "Spatial Scale and Neighbourhood Regeneration in England: A Case Study of Avon." *Environment and Planning C*, 21(5), 651-662.
- Hassan, Y., and Easa, S. M. (2003). "Effect of Vertical Alignment on Driver Perception of Horizontal Curves." *Journal of Transportation Engineering*, 129(4), 399-407.
- Hauer, E. (2000). "Safety of Horizontal Curves." available:  
<http://www.roadsafetyresearch.com/>, Toronto.
- Hauer, E. (2004). "Statistical Safety Modeling." University of Toronto, Toronto.
- Haynes, R., Jones, A., Kennedy, V., Harvey, I., and Jewell, T. (2007a). "District variations in road curvature in England and Wales and their association with road traffic crashes." *Environment and Planning A*, 39, 1222-1237.
- Haynes, R., Lake, I., Kingham, S., Sabel, C., Pearce, J., and Barnett, R. (2007b). "Influence of road curvature on fatal crashes in New Zealand." Unpublished Draft.
- Ikeda, T., and Mori, N. "Analysis of Correlation between Roadway Alignment and Traffic Accidents." *3rd International Symposium on Highway Geometric Design*, Chicago.
- Jackett, M. J. "On which roads do accidents occur? A policy for locating advisory speed signs." *IPENZ Annual Conference*.
- Koorey, G. F. "The Use of Road Geometry Data for Highway Applications." *3rd International Symposium on Highway Geometric Design*, Chicago.
- Koorey, G. F., and Tate, F. N. (1997). "Review of Accident Analysis Procedures for Project Evaluation Manual." *Research Report No. 85*, Transfund, Wellington.
- Lamm, R., Choueiri, E. M., and Mailaender, T. (1992). "Traffic safety on two continents—a ten year analysis of human and vehicular involvements." Proc., Strategic Hwy. Res. Program (SHRP) and Traffic Safety on Two Continents.
- Lamm, R., Hayward, J., and Cargin, J. (1986). "Comparison of different procedures for evaluating speed consistency." *Transp. Res. Rec. 1100*, Transportation Research Board, Washington D.C.
- Landcare Research. (2000). "Digital Elevation Model of New Zealand 25m Cellsize - Version 1." Lincoln.
- Leathwick, J. R., Wilson, G., and Stephens, R. T. T. (2002). "Climate Surfaces for New Zealand." Landcare Research Contract Report: LC9792/126.
- LINZ (Land Information New Zealand). (2005). "New Zealand Geographic Place-Names Database." available:  
["http://www.linz.govt.nz/core/placenames/searchplacenames/downloaddataset/index.html".](http://www.linz.govt.nz/core/placenames/searchplacenames/downloaddataset/index.html)
- Local Government Data Unit. (2004). "Indices of Deprivation Briefing Paper." available: [http://www.lgdu-wales.gov.uk/Documents/Project/Deprivation/WDE02000\\_040600\\_WIMD\\_Indices\\_Briefing\\_eng.pdf](http://www.lgdu-wales.gov.uk/Documents/Project/Deprivation/WDE02000_040600_WIMD_Indices_Briefing_eng.pdf), Cardiff, Wales.
- LTNZ (Land Transport New Zealand). (2004). "A New Zealand Guide to the Treatment of Crash Locations." available: <http://www.ltsa.govt.nz/roads/crash-locations/docs/crash-locations.pdf>.

- Lundeen, M. (2005). "Split Lines at Points." Visual Basic Code, available: <http://arcscrips.esri.com/details.asp?dbid=14190>.
- Mahalel, D., and Szternfeld, Z. (1986). "Safety Improvements and Driver Perception." *Accident Analysis and Prevention*, 18(1), 37-42.
- Maher, M. J., and Summersgill, I. (1996). "A Comprehensive Methodology for the Fitting of Predictive Accident Models." *Accident Analysis & Prevention*, 28(3), 281-296.
- Manley, D., Flowerdew, R., and Steel, D. (2006). "Scales, levels and processes: Studying spatial patterns of British census variables." *Computers, Environment and Urban Systems*, 30(2), 143-160.
- McFadden, J., and Elefteriadou, L. (2000). "Evaluating Horizontal Alignment Design Consistency of Two-Lane Rural Highways: Development of New Procedure." *Transportation Research Record 1737*, 9-17, National Research Council, Washington D.C.
- McLean, J. R. (1989). *Two Lane Highway Traffic Operations*, Gordon and Breach Science Publishers, New York.
- Messer, C. J., Mounce, J. M., and Bracket, R. Q. (1981). "Highway Geometric Design Consistency Related to Highway Expectancy." *FHWA/RD/81/035-037*, Federal Highway Administration Department of Transportation, Washington.
- Miaou, S.-P., and Lum, H. (1993). "Modeling vehicle accidents and highway geometric design relationships." *Accident Analysis & Prevention*, 25(6), 689-709.
- Milton, J. C., and Mannering, F. L. (1996). "Relationship between Highway Geometrics, Traffic Related Elements, and Motor Vehicle Accidents." United States.
- Ministry of Transport. (2004). "Preliminary Results of Household Travel Survey 2003 - 2004." available: <http://www.mot.govt.nz/assets/NewPDFs/Travel-Survey-Year12-preliminary-results-for-release.pdf>.
- Ministry of Transport. (2006). "Motor Vehicle Crashes in New Zealand: Section 1 - Historical." available: <http://www.transport.govt.nz/annual-statistics-2005/>.
- Ministry of Transport. (2007). "Crash Analysis System." available: <http://www.ltsa.govt.nz/research/cas/index.html>.
- Minitab Inc. (2003). "MINITAB Statistical Software." Release 14 for Windows, State College, Pennsylvania.
- Montello, D. R., Neil, J. S., and Paul, B. B. (2001). "Scale in Geography." *International Encyclopedia of the Social & Behavioral Sciences*, Pergamon, Oxford, 13501-13504.
- MWH New Zealand Ltd. (2001). "Upper Hutt City Council 24 hour Recalibrated Spatial Traffic Flow Model." Wellington.
- Nicholson, A. J. (1998). "Superelevation, Side Friction and Roadway Consistency." *American Society of Civil Engineers Journal of Transportation Engineering*, 124(5), 411-418.
- Nicholson, A. J. (2006). *ENTR602 Accident Reduction and Prevention Course notes*, University of Canterbury, Christchurch.
- Nicholson, A. J., and Saleh, M. F. (2004). *Transportation Engineering ENCI 382 Course Notes*, University of Canterbury, Christchurch.
- Noland, R. B., and Oh, L. (2004). "The Effect of Infrastructure and Demographic Change on Traffic-Related Fatalities and Crashes: A Case Study of Illinois County-Level Data." *Accident Analysis and Prevention*, 36(4), 525-532.



- Noland, R. B., and Quddus, M. A. (2004). "A spatially disaggregate analysis of road casualties in England." *Accident Analysis and Prevention*, 36(6), 973-984.
- Park, Y.-J., and Saccomanno, F. F. (2006). "Evaluating speed consistency between successive elements of a two-lane rural highway." *Transportation Research Part A: Policy and Practice*, 40(5), 375-385.
- Petrakis, M., Psiloglou, B., Lianou, M., Keramitsoglou, I., and Cartalis, C. (2005). "Evaluation of forest fire risk and fire extinction difficulty at the mountainous park of Vikos-Aoos, Northern Greece: use of remote sensing and GIS techniques." *International Journal of Risk Assessment and Management*, 5(1), 50.
- Polus, A., and Mattar-Habib, C. (2004). "New Consistency Model for Rural Highways and Its Relationship to Safety." *Journal of Transportation Engineering*, 130(3), 286-293.
- Progressive Insurance. (2002). "Claims Survey Finds There's No Place Like Home - For Vehicle Crashes." [http://newsroom.progressive.com/Progressive\\_insurance\\_news/2002/May/five\\_miles.aspx](http://newsroom.progressive.com/Progressive_insurance_news/2002/May/five_miles.aspx).
- Road Research Laboratory. (1965). "Research on Road Traffic." HMSO, London.
- Shankar, V., Mannering, F., and Barfield, W. (1996). "Statistical analysis of accident severity on rural freeways." *Accident Analysis and Prevention*, 28(3), 391-401.
- Siddall, W. (2005). "Remove Duplicates." Visual Basic Code, available: <http://forums.esri.com/Thread.asp?c=93&f=987&t=164373&mc=9#msgid482599>.
- Statistics New Zealand. (2002). "2001 Census Meshblock Database."
- Statistics New Zealand. (2006). "Explanatory Notes." available: <http://www.stats.govt.nz/urban-rural-profiles/explanatory-notes/default.htm>.
- Statistics New Zealand. (2007). "NZ Population Clock." <http://www.stats.govt.nz>, ed.
- Stephens, S. E. (2004). "The influence of landscape characteristics on duck nesting success in the Missouri Coteau region of North Dakota," Ph.D., Montana State University, United States -- Montana.
- Steven, S. W., Charles, R. B., Jr., Carmen, M. B., Jonathan, A. J., and Chad, J. K. (2004). "Fish-habitat modeling for gap analysis to conserve the endangered Topeka shiner (*Notropis topeka*)." *Canadian Journal of Fisheries and Aquatic Sciences*, 61(6), 954.
- Stutts, J., Knipling, R. R., Pfefer, R., Newman, T. R., Slack, K. L., and Hardy, K. K. (2005). "Guidance for Implementation of the AASHTO Strategic Highway Safety Plan, Volume 14: A Guide for Reducing Crashes Involving Drowsy and Distracted Drivers." *NCHRP Report No. 500*, Transportation Research Board, Washington D.C.
- Torbic, D. J., Harwood, D. W., Gilmore, D. K., Pfefer, R., Newman, T. R., Slack, K. L., and Hardy, K. K. (2004). "Guidance for Implementation of the AASHTO Strategic Highway Safety Plan, Volume 7: A Guide for Reducing Collisions on Horizontal Curves." *NCHRP Report No. 500*, Transportation Research Board, Washington D.C.
- Townsend, P., Phillimore, P., and Beattie, A. (1988). *Health and Deprivation: Inequality and the North*, Croom Helm, London.
- Transit New Zealand. (2006). "State Highway Spatial Traffic Flow Data and Count Site Locations 2001-2006." Wellington.
- Transit New Zealand. (2007). "Frequently Asked Questions." available: <http://www.transit.govt.nz/about/faqs>.

- Vest, A., and Stamaiadis, N. "Use of Warning Signs and Markings to Reduce Speeds on Horizontal Curves." *Third International Symposium on Highway Geometric Design*, Chicago.
- Vis, M. (2007). "European Road Safety Observatory - Horizontal Alignment." available:  
[http://www.erso.eu/knowledge/content/15\\_road/getting\\_initial\\_safety\\_design\\_principles\\_right.htm](http://www.erso.eu/knowledge/content/15_road/getting_initial_safety_design_principles_right.htm).
- Wallis, R. C. (2005). "A GIS model for predicting potential "high risk" areas of West Nile virus by identifying ideal mosquito breeding habitats," M.S., Mississippi State University, United States -- Mississippi.
- Wilson, T. D. (1968). "Road Safety By Design." *J. Inst. of Hwy. Engrs.*, 15, 23-33.
- Wood, G. (2003). "Modelling the Ecological Footprint of Green Travel Plans using GIS and Network Analysis: from Metaphor to Management Tool?" *Environment and Planning B*, 30(4), 523-40.

## 11.0 Appendix

### 11.1 Appendix 1: Summary of Crash Data

Table 11.1 Most Common Movement Types for NZ Fatal Crashes 1996-2005

<b>Crash Code</b>	<b>Description</b>	<b>Total Crashes</b>	<b>Percentage of Total</b>
DA	Lost control turning right	555	13.7
DB	Lost control turning left	464	11.4
BF	Head on collision due to loss of control on corner	370	9.1
CB	Loss of control, off roadway to left	229	5.6
BA	Head on collision on straight	223	5.5
BC	Head on collision swinging wide	219	5.4
NA	Hit pedestrian crossing from left	185	4.6
HA	Crossing at right angle	160	3.9
CC	Loss of control, off roadway to right	159	3.9
AB	Head on collision during overtaking	125	3.1
JA	Crossing collision, vehicle turning right	121	3.0
BE	Head on collision due to loss of control on straight	120	3.0
NB	Hit pedestrian crossing from right	113	2.8
LB	Vehicle making right turn against	97	2.4
BB	Head on collision while cutting corner	88	2.2
All Other Types		828	20.4

## 11.2 Appendix 2: Model Correlations

### 11.2.1 Rural State Highways

Table 11.2 Pearson Product Moment Correlation Coefficients and p-values (*italicised*) for Rural State Highway Variables, part A

	CRASH	Bend Density	Detour Ratio	Cumulative Angle	Mean Angle	Std. Dev. of Angles	Mean Rainfall	Elevation	Std. Dev. of Rainfall
Bend Density	-0.226 <i>0.000</i>								
Detour Ratio	-0.120 <i>0.000</i>	0.183 <i>0.000</i>							
Cumulative Angle	-0.260 <i>0.000</i>	0.593 <i>0.000</i>	0.519 <i>0.000</i>						
Mean Angle	0.031 <i>0.194</i>	-0.373 <i>0.000</i>	0.213 <i>0.000</i>	0.236 <i>0.000</i>					
Std. Dev. of Angles	0.102 <i>0.000</i>	-0.330 <i>0.000</i>	0.091 <i>0.000</i>	0.144 <i>0.000</i>	0.727 <i>0.000</i>				
Mean Rainfall	-0.110 <i>0.000</i>	0.114 <i>0.000</i>	0.124 <i>0.000</i>	0.203 <i>0.000</i>	0.086 <i>0.000</i>	-0.061 <i>0.011</i>			
Elevation	-0.156 <i>0.000</i>	0.129 <i>0.000</i>	0.033 <i>0.176</i>	0.060 <i>0.012</i>	-0.028 <i>0.245</i>	-0.168 <i>0.000</i>	0.180 <i>0.000</i>		
Std. Dev. of Rainfall	0.008 <i>0.737</i>	0.029 <i>0.234</i>	-0.020 <i>0.411</i>	0.052 <i>0.030</i>	0.033 <i>0.166</i>	-0.026 <i>0.279</i>	-0.051 <i>0.034</i>	-0.129 <i>0.000</i>	
Std. Dev. of Elevations	-0.140 <i>0.000</i>	0.243 <i>0.000</i>	0.347 <i>0.000</i>	0.401 <i>0.000</i>	0.038 <i>0.119</i>	-0.109 <i>0.000</i>	0.181 <i>0.000</i>	0.296 <i>0.000</i>	0.013 <i>0.592</i>
Intersection within 30m?	0.075 <i>0.002</i>	-0.040 <i>0.096</i>	0.015 <i>0.523</i>	-0.012 <i>0.633</i>	0.047 <i>0.051</i>	0.130 <i>0.000</i>	-0.067 <i>0.005</i>	-0.083 <i>0.001</i>	-0.052 <i>0.030</i>
Junction Density	0.107 <i>0.000</i>	0.007 <i>0.774</i>	-0.088 <i>0.000</i>	0.105 <i>0.000</i>	0.219 <i>0.000</i>	0.538 <i>0.000</i>	-0.169 <i>0.000</i>	-0.205 <i>0.000</i>	-0.052 <i>0.031</i>
Weighted Junction Density	0.116 <i>0.000</i>	-0.065 <i>0.007</i>	-0.105 <i>0.000</i>	-0.038 <i>0.118</i>	0.142 <i>0.000</i>	0.380 <i>0.000</i>	-0.179 <i>0.000</i>	-0.219 <i>0.000</i>	-0.035 <i>0.149</i>
Population	0.003 <i>0.903</i>	0.095 <i>0.000</i>	0.027 <i>0.264</i>	0.073 <i>0.003</i>	-0.006 <i>0.794</i>	0.037 <i>0.129</i>	0.047 <i>0.050</i>	-0.175 <i>0.000</i>	0.106 <i>0.000</i>
Employment	0.044 <i>0.070</i>	-0.013 <i>0.585</i>	-0.025 <i>0.291</i>	-0.043 <i>0.076</i>	-0.020 <i>0.406</i>	0.025 <i>0.294</i>	-0.020 <i>0.417</i>	-0.070 <i>0.004</i>	0.005 <i>0.836</i>
Meshblock Area	-0.096 <i>0.000</i>	-0.067 <i>0.006</i>	0.037 <i>0.120</i>	-0.003 <i>0.901</i>	0.070 <i>0.004</i>	-0.101 <i>0.000</i>	0.349 <i>0.000</i>	0.321 <i>0.000</i>	0.030 <i>0.209</i>
(Pop + Emp)/MBArea	-0.001 <i>0.972</i>	0.005 <i>0.840</i>	-0.021 <i>0.391</i>	0.046 <i>0.054</i>	0.103 <i>0.000</i>	0.198 <i>0.000</i>	0.024 <i>0.324</i>	-0.053 <i>0.028</i>	0.030 <i>0.219</i>
AADT	0.248 <i>0.000</i>	0.094 <i>0.000</i>	-0.109 <i>0.000</i>	-0.095 <i>0.000</i>	-0.172 <i>0.000</i>	0.015 <i>0.524</i>	-0.166 <i>0.000</i>	-0.272 <i>0.000</i>	0.047 <i>0.052</i>

**Table 11.3 Pearson Product Moment Correlation Coefficients and p-values (*italicised*) for Rural State Highway Variables, part B**

	Std. Dev. of Elevations	Intersection within 30m?	Junction Density	Weighted Junction Density	Population	Employment	Meshblock Area	(Pop + Emp)/MBArea
<b>Intersection within 30m?</b>	-0.072							
	<i>0.003</i>							
<b>Junction Density</b>	-0.173	0.199						
	<i>0.000</i>	<i>0.000</i>						
<b>Weighted Junction Density</b>	-0.206	0.091	0.501					
	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>					
<b>Population</b>	0.021	0.007	0.076	0.082				
	<i>0.373</i>	<i>0.759</i>	<i>0.002</i>	<i>0.001</i>				
<b>Employment</b>	-0.054	0.003	0.046	0.084	0.472			
	<i>0.025</i>	<i>0.914</i>	<i>0.058</i>	<i>0.000</i>	<i>0.000</i>			
<b>Meshblock Area</b>	0.103	-0.079	-0.241	-0.256	-0.030	-0.006		
	<i>0.000</i>	<i>0.001</i>	<i>0.000</i>	<i>0.000</i>	<i>0.208</i>	<i>0.816</i>		
<b>(Pop + Emp)/MBArea</b>	-0.026	0.062	0.339	0.041	0.071	0.102	-0.087	
	<i>0.275</i>	<i>0.010</i>	<i>0.000</i>	<i>0.091</i>	<i>0.003</i>	<i>0.000</i>	<i>0.000</i>	
<b>AADT</b>	-0.124	0.061	0.260	0.151	0.108	0.063	-0.281	0.030
	<i>0.000</i>	<i>0.011</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.009</i>	<i>0.000</i>	<i>0.207</i>

### 11.2.2 Rural non-State Highways

**Table 11.4 Pearson Product Moment Correlation Coefficients and p-values (*italicised*) for Rural non-State Highway Variables, part A**

	<b>CRASH</b>	<b>Bend Density</b>	<b>Detour Ratio</b>	<b>Cumulative Angle</b>	<b>Mean Angle</b>	<b>Std. Dev. of Angles</b>	<b>Mean Rainfall</b>	<b>Elevation</b>
<b>Bend Density</b>	-0.353							
	<i>0.000</i>							
<b>Detour Ratio</b>	-0.130	0.319						
	<i>0.000</i>	<i>0.000</i>						
<b>Cumulative Angle</b>	-0.350	0.724	0.513					
	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>					
<b>Mean Angle</b>	0.040	-0.296	0.221	0.255				
	<i>0.047</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>				
<b>Std. Dev. of Angles</b>	0.112	-0.350	0.104	0.081	0.748			
	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>			
<b>Mean Rainfall</b>	-0.079	0.185	0.105	0.255	0.076	-0.031		
	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.125</i>		
<b>Elevation</b>	-0.190	0.147	0.161	0.202	0.090	-0.051	0.168	
	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.012</i>	<i>0.000</i>	
<b>Std. Dev. of Rainfall</b>	-0.049	0.041	0.053	0.129	0.062	0.071	-0.098	-0.132
	<i>0.014</i>	<i>0.044</i>	<i>0.009</i>	<i>0.000</i>	<i>0.002</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>
<b>Std. Dev. of Elevations</b>	-0.251	0.391	0.373	0.542	0.130	-0.002	0.162	0.340
	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.927</i>	<i>0.000</i>	<i>0.000</i>
<b>Intersection within 30m?</b>	0.212	-0.128	-0.074	-0.139	0.013	0.103	-0.118	-0.096
	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.527</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>
<b>Junction Density</b>	0.103	-0.138	-0.147	-0.124	0.058	0.373	-0.150	-0.211
	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.004</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>
<b>Weighted Junction Density</b>	0.121	-0.173	-0.157	-0.209	-0.002	0.268	-0.116	-0.199
	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.923</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>
<b>Population</b>	0.062	-0.012	-0.015	-0.024	-0.005	0.057	0.025	-0.193
	<i>0.002</i>	<i>0.551</i>	<i>0.457</i>	<i>0.236</i>	<i>0.821</i>	<i>0.005</i>	<i>0.212</i>	<i>0.000</i>
<b>Employment</b>	0.068	-0.056	-0.013	-0.066	-0.033	0.002	0.023	-0.033
	<i>0.001</i>	<i>0.005</i>	<i>0.525</i>	<i>0.001</i>	<i>0.101</i>	<i>0.938</i>	<i>0.246</i>	<i>0.105</i>
<b>Meshblock Area</b>	-0.169	0.076	0.142	0.138	0.069	-0.056	0.127	0.352
	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.001</i>	<i>0.005</i>	<i>0.000</i>	<i>0.000</i>
<b>(Pop + Emp)/MBArea</b>	0.047	-0.064	-0.046	-0.023	0.068	0.169	-0.058	-0.097
	<i>0.019</i>	<i>0.001</i>	<i>0.021</i>	<i>0.263</i>	<i>0.001</i>	<i>0.000</i>	<i>0.004</i>	<i>0.000</i>

**Table 11.5 Pearson Product Moment Correlation Coefficients and p-values (*italicised*) for Rural non-State Highway Variables, part B**

	Std. Dev. of Rainfall	Std. Dev. of Elevations	Intersection within 30m?	Junction Density	Weighted Junction Density	Population	Employment	Meshblock Area
Std. Dev. of Elevations	0.062							
	<i>0.002</i>							
Intersection within 30m?	0.011	-0.130						
	<i>0.596</i>	<i>0.000</i>						
Junction Density	0.156	-0.159	0.314					
	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>					
Weighted Junction Density	-0.002	-0.172	0.142	0.444				
	<i>0.935</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>				
Population	0.150	-0.007	0.044	0.128	0.120			
	<i>0.000</i>	<i>0.735</i>	<i>0.028</i>	<i>0.000</i>	<i>0.000</i>			
Employment	0.005	-0.046	0.022	-0.017	0.075	0.405		
	<i>0.793</i>	<i>0.023</i>	<i>0.272</i>	<i>0.395</i>	<i>0.000</i>	<i>0.000</i>		
Meshblock Area	-0.038	0.175	-0.113	-0.221	-0.200	-0.082	0.007	
	<i>0.060</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.725</i>	
(Pop + Emp)/MBArea	0.165	-0.063	0.147	0.496	0.044	0.170	0.009	-0.115
	<i>0.000</i>	<i>0.002</i>	<i>0.000</i>	<i>0.000</i>	<i>0.027</i>	<i>0.000</i>	<i>0.645</i>	<i>0.000</i>

### 11.2.3 Urban Roads

**Table 11.6 Pearson Product Moment Correlation Coefficients and p-values (*italicised*) for Urban Road Variables, part A**

	<b>CRASH</b>	<b>Bend Density</b>	<b>Detour Ratio</b>	<b>Cumulative Angle</b>	<b>Mean Angle</b>	<b>Std. Dev. of Angles</b>	<b>Mean Rainfall</b>	<b>Elevation</b>
<b>Bend Density</b>	-0.096							
	<i>0.000</i>							
<b>Detour Ratio</b>	-0.096	1.000						
	<i>0.000</i>	<i>0.000</i>						
<b>Cumulative Angle</b>	-0.097	1.000	1.000					
	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>					
<b>Mean Angle</b>	-0.080	0.995	0.995	0.995				
	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>				
<b>Std. Dev. of Angles</b>	-0.072	0.995	0.995	0.995	0.997			
	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>			
<b>Mean Rainfall</b>	0.028	-0.041	-0.041	-0.041	-0.046	-0.047		
	<i>0.087</i>	<i>0.010</i>	<i>0.010</i>	<i>0.010</i>	<i>0.005</i>	<i>0.003</i>		
<b>Elevation</b>	-0.112	0.010	0.010	0.010	0.009	0.000	0.210	
	<i>0.000</i>	<i>0.543</i>	<i>0.540</i>	<i>0.533</i>	<i>0.579</i>	<i>0.985</i>	<i>0.000</i>	
<b>Std. Dev. of Rainfall</b>	-0.077	0.036	0.036	0.036	0.034	0.031	-0.266	-0.153
	<i>0.000</i>	<i>0.028</i>	<i>0.027</i>	<i>0.027</i>	<i>0.037</i>	<i>0.054</i>	<i>0.000</i>	<i>0.000</i>
<b>Std. Dev. of Elevations</b>	-0.239	0.003	0.003	0.004	-0.008	-0.018	0.070	0.295
	<i>0.000</i>	<i>0.849</i>	<i>0.835</i>	<i>0.805</i>	<i>0.603</i>	<i>0.257</i>	<i>0.000</i>	<i>0.000</i>
<b>Intersection within 30m?</b>	0.204	-0.045	-0.045	-0.046	-0.024	-0.017	-0.055	-0.061
	<i>0.000</i>	<i>0.005</i>	<i>0.005</i>	<i>0.005</i>	<i>0.137</i>	<i>0.298</i>	<i>0.001</i>	<i>0.000</i>
<b>Junction Density</b>	-0.085	-0.042	-0.042	-0.042	-0.008	0.008	-0.037	-0.103
	<i>0.000</i>	<i>0.009</i>	<i>0.009</i>	<i>0.009</i>	<i>0.611</i>	<i>0.621</i>	<i>0.021</i>	<i>0.000</i>
<b>Weighted Junction Density</b>	-0.144	0.063	0.063	0.063	0.038	0.026	0.064	0.078
	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.020</i>	<i>0.110</i>	<i>0.000</i>	<i>0.000</i>
<b>Population</b>	-0.127	0.006	0.005	0.006	-0.002	-0.002	0.057	-0.067
	<i>0.000</i>	<i>0.732</i>	<i>0.734</i>	<i>0.727</i>	<i>0.909</i>	<i>0.904</i>	<i>0.000</i>	<i>0.000</i>
<b>Employment</b>	0.131	-0.033	-0.033	-0.033	-0.031	-0.027	-0.001	-0.074
	<i>0.000</i>	<i>0.041</i>	<i>0.041</i>	<i>0.041</i>	<i>0.053</i>	<i>0.092</i>	<i>0.951</i>	<i>0.000</i>
<b>Meshblock Area</b>	-0.004	0.011	0.011	0.011	0.003	-0.002	-0.021	0.016
	<i>0.787</i>	<i>0.513</i>	<i>0.512</i>	<i>0.515</i>	<i>0.877</i>	<i>0.926</i>	<i>0.200</i>	<i>0.325</i>
<b>(Pop + Emp)/MBArea</b>	0.067	-0.117	-0.117	-0.117	-0.092	-0.084	-0.015	-0.087
	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.350</i>	<i>0.000</i>



**Table 11.7 Pearson Product Moment Correlation Coefficients and p-values (*italicised*) for Urban Road Variables, part B**

	Std. Dev. of Rainfall	Std. Dev. of Elevations	Intersection within 30m?	Junction Density	Weighted Junction Density	Population	Employment	Meshblock Area
Std. Dev. of Elevations	0.136							
	<i>0.000</i>							
Intersection within 30m?	-0.021	-0.073						
	<i>0.200</i>	<i>0.000</i>						
Junction Density	0.006	-0.038	0.241					
	<i>0.718</i>	<i>0.020</i>	<i>0.000</i>					
Weighted Junction Density	0.023	0.073	-0.273	-0.515				
	<i>0.157</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>				
Population	0.025	0.024	-0.032	0.072	0.026			
	<i>0.119</i>	<i>0.142</i>	<i>0.045</i>	<i>0.000</i>	<i>0.104</i>			
Employment	-0.004	-0.084	0.079	0.004	-0.039	-0.077		
	<i>0.797</i>	<i>0.000</i>	<i>0.000</i>	<i>0.804</i>	<i>0.016</i>	<i>0.000</i>		
Meshblock Area	0.025	0.008	-0.049	-0.186	0.132	0.007	0.001	
	<i>0.116</i>	<i>0.619</i>	<i>0.002</i>	<i>0.000</i>	<i>0.000</i>	<i>0.681</i>	<i>0.953</i>	
(Pop + Emp)/MBArea	0.033	0.008	0.181	0.486	-0.337	0.056	0.178	-0.094
	<i>0.041</i>	<i>0.625</i>	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.001</i>	<i>0.000</i>	<i>0.000</i>